

STIC-Biotech/ChemLib

191514

From: Chan, Christina
Sent: Wednesday, May 31, 2006 1:22 PM
To: Marvich, Maria; STIC-Biotech/ChemLib
Subject: RE: 09/729658 SEQ ID NO:2

~~Please rush.~~ Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841.
Remsen, 3E89

CRFB

-----Original Message-----

From: Marvich, Maria
Sent: Wednesday, May 31, 2006 1:21 PM
To: Chan, Christina
Subject: FW: 09/729658 SEQ ID NO:2

29301

Please rush an after-final sequence search that is otherwise in condition for allowance.

Please include interference.

Thank you.

Maria Bonovich Marvich, PhD
United States Patent and Trademark Office
Remsen 2B84
AU 1633
Mail Box 2C70
571-272-0774

Searcher: Thib
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: 6/8/06
Searcher Prep Time: 5
Online Time: 5

Type of Search /
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Compugen
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:21:28 ; Search time 200 Seconds
(without alignments)
893.858 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYEVEVERELLPAAAPRR.....NMSKHTTFPGAIRLGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq.8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2098	100.0	391	5	ABP53104 Human ect
2	2098	100.0	391	5	ABP53104 Human ect
3	2098	100.0	391	5	ABP53104 Human ect
4	2098	100.0	391	7	ABP53104 Human ect
5	2098	100.0	391	7	ABP53104 Human ect
6	2098	100.0	391	7	ABP53104 Human ect
7	2098	100.0	391	9	ABP53104 Human ect
8	2098	100.0	391	9	ABP53104 Human ect
9	1965	93.7	391	9	ABP53104 Human ect
10	1873	89.3	377	5	ABP53104 Human ect
11	1873	89.3	377	7	ABP53104 Human ect
12	779	37.1	412	7	ADP75166 Fusion co
13	759	36.2	410	7	ADP75166 Fusion co
14	700	33.4	135	2	AAW44764 Human anh
15	294	14.0	357	2	AAW57646 Collagen
16	294	14.0	357	2	AAW57646 Collagen
17	294	14.0	357	4	AAW57646 Collagen
18	294	14.0	357	4	AAW57646 Collagen
19	286	13.6	742	8	ADL91066 Mouse col
20	282	13.4	742	4	ADL91066 Mouse col
21	282	13.4	742	8	ADL91066 Mouse col
22	282	13.4	742	8	ADL91066 Mouse col
23	280.5	13.4	1838	2	AAW53257 Human col

24	280.5	13.4	1838	7	ADP55566 Human Pro
25	280.5	13.4	1838	7	ADP55566 Human Pro
26	280.5	13.4	1838	7	ADP55566 Human Pro
27	280.5	13.4	1838	9	ADP55566 Human Pro
28	280.5	13.4	1838	9	ADP55566 Human Pro
29	280.5	13.4	1838	9	ADP55566 Human Pro
30	280.5	13.4	1838	9	ADP55566 Human Pro
31	277	13.2	330	2	AAW57645 Collagen
32	277	13.2	408	2	AAW57645 Collagen
33	277	13.2	408	4	AAW57645 Collagen
34	277	13.2	408	4	AAW57645 Collagen
35	276	13.2	622	8	ADL91065 Human col
36	276	13.2	622	8	ADL91065 Human col
37	275.5	13.1	180	9	ADP56200 Human PRO
38	275.5	13.1	180	9	ADP56200 Human PRO
39	275.5	13.1	180	9	ADP56200 Human PRO
40	275.5	13.1	180	10	ADP56200 Human PRO
41	275	13.1	1466	4	AAE02534 Bovine al
42	275	13.1	1466	4	AAE02534 Bovine al
43	273.5	13.0	1014	9	ADW99577 Human gel
44	273	13.0	342	4	AAW39686 Human pol
45	273	13.0	547	3	AAW77985 Human col

ALIGNMENTS

RESULT 1	ABP53104	ABP53104 standard; protein; 391 AA.
ID	ABP53104	ABP53104 standard; protein; 391 AA.
XX	ABP53104	ABP53104 standard; protein; 391 AA.
AC	ABP53104	ABP53104 standard; protein; 391 AA.
XX	ABP53104	ABP53104 standard; protein; 391 AA.
DT	11-NOV-2002	(first entry)
XX	11-NOV-2002	(first entry)
DE	Human ectodermal dysplasia protein (EDA) SEQ ID NO:6.	
XX	Human ectodermal dysplasia protein (EDA) SEQ ID NO:6.	
KW	Human; adipocyte complement related protein; zacrpx2; antiinflammatory; vasotropic; antimicrobial; vulnary; antiarteriosclerotic; gene therapy; adipogenesis; gluconeogenesis; lipogenesis; ischaemia; glucose uptake; protein synthesis; thermogenesis; oxygen utilisation; inflammation; neurotransmitter; antimicrobial; infection.	
OS	Homo sapiens.	
XX	Homo sapiens.	
PN	WO200246417-A2.	
XX	WO200246417-A2.	
PD	13-JUN-2002.	
XX	13-JUN-2002.	
PF	07-DEC-2001; 2001WO-US047348.	
XX	07-DEC-2001; 2001WO-US047348.	
PR	07-DEC-2000; 2000US-0254019P.	
XX	07-DEC-2000; 2000US-0254019P.	
PA	(ZYMO) ZYMOGENETICS INC.	
XX	(ZYMO) ZYMOGENETICS INC.	
PI	Haldeman BA, Thayer EC, Sheppard PO;	
XX	Haldeman BA, Thayer EC, Sheppard PO;	
DR	WPI; 2002-657398/70.	
XX	WPI; 2002-657398/70.	
PT	Novel human adipocyte complement related protein, zacrpx2, useful for modulating energy balance in mammals, and preventing injury due to ischemia/inflammation.	
XX	Novel human adipocyte complement related protein, zacrpx2, useful for modulating energy balance in mammals, and preventing injury due to ischemia/inflammation.	
PS	Disclosure; Page 111-112; 118pp; English.	
XX	Disclosure; Page 111-112; 118pp; English.	
CC	The present invention describes a human adipocyte complement related protein (I) designated zacrpx2. (I) has antiinflammatory, vasotropic, antimicrobial, vulnary and antiarteriosclerotic activities, and can be used in somatic cell gene therapy, and as a modulator of neurotransmission. (I) is useful for modulating energy balance in mammals or for protecting endothelial cells from injury, for modulating cellular metabolic reactions e.g. adipogenesis, gluconeogenesis, lipogenesis, and lipogenesis, glucose uptake, protein synthesis, thermogenesis, and oxygen	

XX 11-AUG-2003 (first entry)
XX Human EDA protein.
DE Human; EDA; tumour necrosis factor; ligand; cytostatic; immunomodulator;
KW osteopathic.
XX Homo sapiens.
OS WO2003040307-A2.
PN 15-MAY-2003.
XX 25-JUL-2002; 2002WO-US023782.
XX 27-JUL-2001; 2001US-0307838P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Hilbert DH, Rosen CA;
XX WPI; 2003-430659/40.
DR N-PSDB; ACC57910.
XX New heteromultimeric complex having a first polypeptide member of the
PT tumor necrosis factor (TNF) ligand family, and a second different member
PT of TNF ligand family, useful for treating cancer, osteoporosis or an
PT autoimmune disease.
XX Disclosure; Page 387-388; 388pp; English.
XX The present sequence is the protein sequence of human EDA polypeptide.
CC The invention relates to compositions comprising heterotrimeric complexes
CC of tumor necrosis factor (TNF) ligand family members, and their use in
CC the detection, prevention and treatment of disease. In preferred
CC embodiments, the heterotrimeric complex comprises full-length or
CC extracellular portions of EDA together with full-length or extracellular
CC portions of other TNF ligand family members. The heterotrimeric complexes
CC of the invention are useful for treating an autoimmune disease, cancer or
CC osteoporosis, and particularly for inhibiting cancer cell proliferation,
CC increasing B cell proliferation, or inducing apoptosis of T cells
XX
SQ Sequence 391 AA;
Query Match 100.0%; Score 2098; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGFFGLSLALHLLTLCC 60
DB 1 MGYPERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGFFGLSLALHLLTLCC 60
QY 61 YLELSERRRERGAERLGGSTGTSGLSLGLDPSBITSHLQSPSPKQPLEPGE 120
DB 61 YLELSERRRERGAERLGGSTGTSGLSLGLDPSBITSHLQSPSPKQPLEPGE 120
QY 121 AALHSDSDQHOMALLNFFPDEKPYSEESRVRNRKSKNEGADGPKVKKKGKXAG 180
DB 121 AALHSDSDQHOMALLNFFPDEKPYSEESRVRNRKSKNEGADGPKVKKKGKXAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQKNDLSGGLVNDWSRITNPKVFKLHPRSGEVLVDGT 300
DB 241 AGTRENQPAVHLQGGSAIQKNDLSGGLVNDWSRITNPKVFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYVYINFTDFASVEVVDKPELPQCTRSIETGKTNNYNTCYTAGVCLLKARQKI 360
DB 301 YFIYSQVEVYVYINFTDFASVEVVDKPELPQCTRSIETGKTNNYNTCYTAGVCLLKARQKI 360

QY 361 AVKMVHADISINNSKHTTFFGAIIRLGEAPAS 391
DB 361 AVKMVHADISINNSKHTTFFGAIIRLGEAPAS 391
RESULT 4
AB001928
ID ABO01928 standard; protein; 391 AA.
XX ABO01928;
XX 15-AUG-2003 (first entry)
XX Human EDAL-II protein.
XX Human; EDAL-II; hair follicle development; tooth development;
KW sweat gland development; ectodermal disease; alopecia; skin trauma; burn;
KW X-linked hypohidrotic ectodermal dysplasia; XLHED; breast cancer; HED;
KW autosomal recessive hypohidrotic ectodermal dysplasia; hirsutism;
KW ectopic teeth.
XX Homo sapiens.
XX US2003023991-A1.
XX 30-JAN-2003.
XX 04-DEC-2000; 2000US-00729659.
XX 09-JUL-1998; 98US-0092279P.
PR 15-DEC-1998; 98US-0112366P.
PR 29-JUN-1999; 99US-00342681.
XX (ZONA/) ZONANA J.
PA (FERG/) FERGUSON B M.
PA (HEAD/) HEADON D.
PA (OVER/) OVERBEEK P.
XX Zonana J, Ferguson BM, Headon D, Overbeek P;
PI WPI; 2003-456312/43.
DR N-PSDB; ACD07895.
XX Increasing or decreasing hair follicle development, tooth development, or
PT sweat gland development, in a tissue, involves modulating ectodermal
PT dysplasia protein isoform II activity in the tissue.
XX Claim 26; Fig 1; 102pp; English.
XX The invention relates to a method of increasing or decreasing one or more
CC of hair follicle development, tooth development, or sweat gland
CC development, in a tissue, involves altering ectodermal dysplasia (EDAL)
CC isoform II (EDAL-II) activity in the tissue. The method is useful for
CC increasing or decreasing hair follicle development, tooth development or
CC sweat gland development, in a tissue of a patient suffering from an
CC ectodermal disease such as X-linked hypohidrotic ectodermal dysplasia
CC (XLHED), autosomal recessive hypohidrotic ectodermal dysplasia (HED) or
CC alopecia. The method is useful for decreasing hair follicle development
CC in a subject suffering from hirsutism. The method is also useful for
CC treating breast cancer and ectopic teeth, burns and trauma of skin due to
CC surgery. The present sequence represents the amino acid sequence of human
CC EDAL-II
XX Sequence 391 AA;
Query Match 100.0%; Score 2098; DB 7; Length 391;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGFFGLSLALHLLTLCC 60
DB 1 MGYPERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGFFGLSLALHLLTLCC 60

PR 27-APR-2000; 2000US-00559290.
 PR 07-JUL-2000; 2000US-0216879P.
 PR 26-MAR-2001; 2001US-0278449P.
 PR 06-JUL-2001; 2001US-00899059.
 PR 24-AUG-2001; 2001US-0314381P.
 PR 07-DEC-2001; 2001US-0336695P.
 PR 23-AUG-2002; 2002US-00226294.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Yu G, Ni J, Rosen CA, Zhang J, Wei P;
 DR WPI; 2003-852773/79.
 DR N-PSDB; AAD63922.
 XX
 PT Use of tumor necrosis factor gamma-beta antagonists for treating or
 PT ameliorating a disease or disorders of the gastrointestinal tract, e.g.
 PT inflammatory bowel disease, Crohn's disease or ulcerative colitis.
 XX
 PS Disclosure; Page 156-157; Opp; English.
 CC The invention relates to methods and compositions for treating or
 CC ameliorating a disease or disorder of the gastrointestinal tract. The
 CC method involves administering a composition comprising tumour necrosis
 CC factor (TNF)-gamma-beta and its receptors DR3 and TR6 to a person with,
 CC or suspected of having the disease or disorder. The antagonist of TNF-
 CC gamma-beta is useful for treating or ameliorating a gastrointestinal
 CC tract disease or disorder, specifically an inflammatory bowel disease,
 CC e.g. Crohn's disease or ulcerative colitis. The present sequence is human
 CC EDA protein. This sequence is used to illustrate the method of the
 CC invention
 XX
 SQ Sequence 391 AA;
 Query Match 100.0%; Score 2098; DB 7; Length 391;
 Best Local Similarity 100.0%; Pred. No. 6.2e-155;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTCC 60
 DB 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTCC 60
 QY 61 YLELSELRERGAESRLGSGTCTGCTSLSLGGLDPSPTSHLGOPSPKQPLEPGE 120
 DB 61 YLELSELRERGAESRLGSGTCTGCTSLSLGGLDPSPTSHLGOPSPKQPLEPGE 120
 QY 121 AALHSDSDQGHQMALNFFPFDEKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 180
 DB 121 AALHSDSDQGHQMALNFFPFDEKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 180
 QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
 DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
 QY 241 AGTRENQAVVHLOGGSAIQVKNLDSGVLDNSRITWNPVKVFKLHRSSELEVLVDGT 300
 DB 241 AGTRENQAVVHLOGGSAIQVKNLDSGVLDNSRITWNPVKVFKLHRSSELEVLVDGT 300
 QY 301 YFIYSQVEVYINFTDFASVEVVDKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 360
 DB 301 YFIYSQVEVYINFTDFASVEVVDKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 360
 QY 361 AVKMVHADISINMSKHTFFFGAIRLGEAPAS 391
 DB 361 AVKMVHADISINMSKHTFFFGAIRLGEAPAS 391
 RESULT 7
 ID ADX06863
 XX ADX06863 standard; protein; 391 AA.
 AC
 XX ADX06863;

DT 21-APR-2005 (first entry)
 DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 1428.
 XX
 KW cytostatic; cyclin-dependent kinase; cdk; biomarker.
 OS Homo sapiens.
 XX
 PN WO2005012875-A2.
 XX
 PD 10-FEB-2005.
 XX
 PF 29-JUL-2004; 2004WO-US024424.
 XX
 PR 29-JUL-2003; 2003US-0490890P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
 XX WPI; 2005-163068/17.
 DR N-PSDB; ADX06862.
 XX
 PT Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.
 XX
 PS Claim 5; SEQ ID NO 1428; 141pp; English.
 CC This invention describes a novel method of predicting or determining
 CC whether a mammal will respond or is responding to an anti-cancer agent
 CC that modulates cyclin-dependent kinase (cdk) activity. The method
 CC comprises measuring the level of one or more biomarkers selected from
 CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
 CC invention is utilized in a kit for determining or predicting whether
 CC patient would be susceptible or resistant to treatment by an agent
 CC modulating cdk activity. The invention also describes a method for
 CC utilizing individualized genetic profiles for treating diseases and
 CC disorders based on patient's response and molecular level, specialized
 CC microarrays comprising the biomarkers described, antibodies directed
 CC against the biomarkers and a cell culture model to identify biomarkers.
 CC The cdk modulator is preferably N-5-[(5-(1,1-Dimethylethyl)-2-
 CC oxazolyl)methyl]thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
 CC tartaric acid salt. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published pct sequences. This
 CC sequence represents a biomarker used in the method of the invention.
 XX
 SQ Sequence 391 AA;
 Query Match 100.0%; Score 2098; DB 9; Length 391;
 Best Local Similarity 100.0%; Pred. No. 6.2e-155;
 Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTCC 60
 DB 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTCC 60
 QY 61 YLELSELRERGAESRLGSGTCTGCTSLSLGGLDPSPTSHLGOPSPKQPLEPGE 120
 DB 61 YLELSELRERGAESRLGSGTCTGCTSLSLGGLDPSPTSHLGOPSPKQPLEPGE 120
 QY 121 AALHSDSDQGHQMALNFFPFDEKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 180
 DB 121 AALHSDSDQGHQMALNFFPFDEKPYSEESRRVRNRKRSKNEGADGPVKNNKKGKAG 180
 QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
 DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
 QY 241 AGTRENQAVVHLOGGSAIQVKNLDSGVLDNSRITWNPVKVFKLHRSSELEVLVDGT 300
 DB 241 AGTRENQAVVHLOGGSAIQVKNLDSGVLDNSRITWNPVKVFKLHRSSELEVLVDGT 300

Db 241 AGTRENQPAVHLQGGSAIQVKNDLSGGVLDMSRITMNPVKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSLETGKTNNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSLETGKTNNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 8
AED55185
ID AED55185 standard; protein; 391 AA.
XX AED55185;
XX
DT 29-DEC-2005 (first entry)
XX
DE Ectodysplasin A protein SEQ ID NO:8.
XX
KW tumour necrosis factor; apoptosis; apoptosis modulation; antibacterial;
KW immunosuppressive; antiinflammatory; cytostatic; anorectic; antidiabetic;
KW antirheumatic; antiarthritic; gastrointestinal-gen.; endotoxigenic shock;
KW infection; inflammation; tumor; neoplasm; insulin resistance;
KW endocrine disease; rheumatoid arthritis; immune disorder;
KW musculoskeletal disease; inflammatory bowel disease;
KW gastrointestinal disease; Osteoprotegeron.
XX
OS Unidentified.
XX
PN US2005227283-A1.
XX
PD 13-OCT-2005.
XX
PF 01-JUN-2005; 2005US-00142736.
XX
PR 21-MAR-2000; 2000US-0190816P.
PR 20-MAR-2001; 2001US-00813329.
XX
PA (CARR/) CARROLL P M.
PA (CHEN/) CHEN J.
PA (RAMA/) RAMANATHAN C S.
PA (XIAO/) XIAO H.
PA (GUAN/) GUAN B.
PA (BOWE/) BOWEN M A.
XX
PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;
XX WPI; 2005-756087/77.
DR
XX
PT New Drosophila tumor necrosis factor polypeptide, useful for treating
PT diseases, e.g. endotoxigenic shock, inflammation, rheumatoid arthritis, or
PT inflammatory bowel disease.
XX
PS Disclosure; SEQ ID NO 8; 117pp; English.
XX
XX The invention relates to a novel isolated Drosophila tumor necrosis
CC factor (DmTNF) polypeptide. An isolated Drosophila tumor necrosis factor
CC (DmTNF) polypeptide comprises a polypeptide sequence selected from: (i)
CC an isolated polypeptide comprising amino acids 1-409, 2-409, 53-409, 62-
CC 409 of a sequence comprising 409 amino acids (AED55183), or (ii) an
CC isolated polypeptide comprising 332 contiguous amino acids of (AED55183),
CC where the polypeptide induces apoptosis in a cell in which the
CC polypeptide is recombinantly expressed, where the isolated polypeptide is
CC produced by a method, which comprises culturing an isolated recombinant
CC host cell comprising a vector comprising the coding region encoding the
CC polypeptide under conditions such that the polypeptide is expressed and
CC recovering the polypeptide. A polypeptide of the invention has
CC antibacterial, immunosuppressive, antiinflammatory, cytostatic,
CC anorectic, antidiabetic, antirheumatic, antiarthritic, and
CC gastrointestinal-gen. activity. The polypeptide, nucleic acids,
CC composition, and methods of the invention are useful for treating

CC diseases, e.g. endotoxigenic shock, inflammation, hemorrhagic necrosis of
CC tumors, cytotoxicity, obesity-linked insulin resistance, rheumatoid
CC arthritis, or inflammatory bowel disease. The present sequence represents
CC an ectodysplasin A protein. Note: The sequence is stated in the
CC disclosure of the invention as being of human origin, but is in the
CC sequence listing as Drosophila melanogaster and has therefore been
CC indexed as unidentified.
XX
SQ Sequence 391 AA;
Query Match 100.0%; Score 2098; DB 9; Length 391;
Best Local Similarity 100.0%; Pred. No. 6.2e-155;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRRGSGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db 1 MGYPEVERRELLPAAAPRRGSGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQPLEPGE 120
QY 121 AALHSDSDQGHOMALLNFFFPDEKPYSEESRVRNRKSKNEGADGVKNKKKKKAG 180
Db 121 AALHSDSDQGHOMALLNFFFPDEKPYSEESRVRNRKSKNEGADGVKNKKKKKAG 180
QY 181 PPGNGPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNGPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQVKNDLSGGVLDMSRITMNPVKLHPRSGEVLVDGT 300
Db 241 AGTRENQPAVHLQGGSAIQVKNDLSGGVLDMSRITMNPVKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSLETGKTNNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYINFTDFASYEVVVDKPFLOCTRSLETGKTNNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 9
AED55186
ID AED55186 standard; protein; 391 AA.
XX AED55186;
XX
DT 29-DEC-2005 (first entry)
XX
DE Murine ectodysplasin A protein SEQ ID NO:9.
XX
KW tumour necrosis factor; apoptosis; apoptosis modulation; antibacterial;
KW immunosuppressive; antiinflammatory; cytostatic; anorectic; antidiabetic;
KW antirheumatic; antiarthritic; gastrointestinal-gen.; endotoxigenic shock;
KW infection; inflammation; tumor; neoplasm; insulin resistance;
KW endocrine disease; rheumatoid arthritis; immune disorder;
KW musculoskeletal disease; inflammatory bowel disease;
KW gastrointestinal disease; Osteoprotegeron.
XX
OS Mus musculus.
XX
PN US2005227283-A1.
XX
PD 13-OCT-2005.
XX
PF 01-JUN-2005; 2005US-00142736.
XX
PR 21-MAR-2000; 2000US-0190816P.
PR 20-MAR-2001; 2001US-00813329.
XX
PA (CARR/) CARROLL P M.

PA	(CHEN/) CHEN J.	AB090983	
PA	(RAMA/) RAMANATHAN C S.	ID	ABB09083 standard; protein; 377 AA.
PA	(XIAO/) XIAO H.	XX	
PA	(GUAN/) GUAN B.	AC	ABB09083;
PA	(BOWE/) BOWEN M A.	XX	
XX		DT	26-JUN-2002 (first entry)
PI	Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;	XX	
XX	WPI; 2005-756087/77.	DE	Mouse Tabby (Ta) protein SEQ ID NO:4.
DR		XX	
XX		KW	Human; mouse; EDAL-II; X-linked hypohidrotic ectodermal dysplasia; XLHED;
XX		KW	DL receptor; DL gene; downless gene; ectodermal dysplasia;
PT	New Drosophila tumor necrosis factor polypeptide, useful for treating	KW	hair follicle growth; tooth growth; endocrine sweat gland development;
PT	diseases, e.g. endotoxemic shock, inflammation, rheumatoid arthritis, or	KW	mammary epithelial tissue growth; breast epithelium.
PT	inflammatory bowel disease.	XX	
XX		OS	Mus musculus.
PS	Disclosure; SEQ ID NO 9; 117pp; English.	PN	
XX		XX	US6355782-B1.
CC	The invention relates to a novel isolated Drosophila tumor necrosis	XX	
CC	factor (DmTNF) polypeptide. An isolated Drosophila tumor necrosis factor	PD	12-MAR-2002.
CC	(DmTNF) polypeptide comprises a polypeptide sequence selected from: (i)	XX	
CC	an isolated polypeptide comprising amino acids 1-409, 2-409, 53-409, 62-	PF	29-JUN-1999; 99US-00342681.
CC	409 of a sequence comprising 409 amino acids (AED55183), or (ii) an	XX	
CC	isolated polypeptide comprising 332 contiguous amino acids of (AED55183),	PR	09-JUL-1998; 98US-0092279P.
CC	where the polypeptide induces apoptosis in a cell in which the	PR	15-DEC-1998; 98US-0112366P.
CC	polypeptide is recombinantly expressed, where the isolated polypeptide is	XX	
CC	produced by a method, which comprises culturing an isolated recombinant	PA	(BAYU) BAYLOR COLLEGE MEDICINE.
CC	host cell comprising a vector comprising the coding region encoding the	PA	(UYOR-) UNIV OREGON HEALTH SCI.
CC	polypeptide under conditions such that the polypeptide is expressed and	XX	
CC	recovering the polypeptide. A polypeptide of the invention has	XX	Zonana J, Ferguson BM, Headon D, Overbeek P;
CC	antibacterial, immunosuppressive, antiinflammatory, cytostatic,	XX	
CC	anorectic, antidiabetic, antirheumatic, antiarthritic, and	XX	
CC	gastrointestinal-gen. activity. The polypeptide, nucleic acids,	DR	WPI; 2002-360478/39.
CC	composition, and methods of the invention are useful for treating	DR	N-PSDB; ABL51010.
CC	diseases, e.g. endotoxemic shock, inflammation, hemorrhagic necrosis of	XX	
CC	tumors, cytotoxicity, obesity-linked insulin resistance, rheumatoid	PT	Novel purified hypohidrotic ectodermal dysplasias protein, useful for
CC	arthritis, or inflammatory bowel disease. The present sequence represents	PT	promoting the development of hair follicles and tooth growth.
CC	murine ectodysplasmin A protein.	XX	
XX		PS	Example 1; Col 71-74; 86pp; English.
XX		XX	
SQ	Sequence 391 AA;	CC	
	Query Match 93.7%; Score 1965; DB 9; Length 391;	CC	The present invention describes a purified hypohidrotic ectodermal
	Best Local Similarity 94.6%; Pred. No. 1.4e-144;	CC	dysplasias protein (I) that promotes the development of hair follicles.
	Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;	CC	(I) is a promoter of hair follicle growth and tooth growth. (I) is useful
		CC	for the development of hair follicles. (I) is also useful for commercial
QY	1 MGYPEVERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60	CC	and clinical diagnostic testing for ectodermal dysplasia (ED). (I) is
Db	1 MGYPEVERREPLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60	CC	also useful as a stimulant for hair and tooth growth, for stimulating
QY	61 YLELSELRRERGAESRLGSGTPTGTSGLSLGGLDPSITSHLGQSPKQPLEPGE 120	CC	endocrine sweat gland development in individuals from whom the normal
Db	61 YLELSELRRERGTESRLGSGPAGTSGTSLSPGSLDPVGPITRHLGQSFQQQPLEPGE 120	CC	sweating mechanism is compromised by disease or surgery, for stimulating
QY	121 AALHSDSQDGHQWALLNFFPFDEKPYSEESRRVRNRKSKNEGADGVPVKNKKGKAG 180	CC	the growth of mammary epithelial tissue, either for reconstructing or
Db	121 DPLPPDSQDRHQWALLNFFPFDEKAYSEESRRVRNRKSKSGEGADGVPVKNKKGKAG 180	CC	cosmetic purposes, and for promoting or maintaining differentiation of
QY	181 PFGPNPQPPGPPGPPGQPGIPGIPGTTVMGPPPPGPPGPPGPPGPPGPPGPPGPPG 240	CC	breast epithelium. (I) is also useful for functional analysis, antibody
Db	181 PFGPNPQPPGPPGPPGQPGIPGIPGTTVMGPPPPGPPGPPGPPGPPGPPGPPGPPG 240	CC	production and patient therapy. ABL51009 to ABL51122 and ABB09082 to
QY	241 ACTRENQRAVHLQGGSAIQKNDLSGVNDWSRITWNPVKVFKLHPRSSELEVLVDGT 300	CC	ABB09090 represent sequences used in the exemplification of the present
Db	241 TGTRENQRAVHLQGGGSAIQKNDLSGVNDWSRITWNPVKVFKLHPRSSELEVLVDGT 300	XX	invention
QY	301 YFIYSQVEVYVYINFTDFASEYVVDSEKPFLOQTRSIETGKTNTYNTCYTAGVCLLKARQKI 360	SQ	Sequence 377 AA;
Db	301 YFIYSQVEVYVYINFTDFASEYVVDSEKPFLOQTRSIETGKTNTYNTCYTAGVCLLKARQKI 360		Query Match 89.3%; Score 1873; DB 5; Length 377;
QY	361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391		Best Local Similarity 90.8%; Pred. No. 2e-137;
Db	361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391		Matches 355; Conservative 2; Mismatches 20; Indels 14; Gaps 1;
		QY	1 MGYPEVERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
		Db	1 MGYPEVERREPLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
		QY	61 YLELSELRRERGAESRLGSGTPTGTSGLSLGGLDPSITSHLGQSPKQPLEPGE 120
		Db	61 YLELSELRRERGTESRLGSGPAGTSGTSLSPGSLDPVGPITRHLGQSFQQQPLEPGE 120
		QY	121 AALHSDSQDGHQWALLNFFPFDEKPYSEESRRVRNRKSKNEGADGVPVKNKKGKAG 180
		Db	121 DPLPPDSQDRHQWALLNFFPFDEKAYSEESRRVRNRKSKSGEGADGVPVKNKKGKAG 180
		QY	181 PFGPNPQPPGPPGPPGQPGIPGIPGTTVMGPPPPGPPGPPGPPGPPGPPGPPGPPG 240
		Db	181 PFGPNPQPPGPPGPPGQPGIPGIPGTTVMGPPPPGPPGPPGPPGPPGPPGPPGPPG 240
		QY	241 ACTRENQRAVHLQGGSAIQKNDLSGVNDWSRITWNPVKVFKLHPRSSELEVLVDGT 300
		Db	241 TGTRENQRAVHLQGGGSAIQKNDLSGVNDWSRITWNPVKVFKLHPRSSELEVLVDGT 300
		QY	301 YFIYSQVEVYVYINFTDFASEYVVDSEKPFLOQTRSIETGKTNTYNTCYTAGVCLLKARQKI 360
		Db	301 YFIYSQVEVYVYINFTDFASEYVVDSEKPFLOQTRSIETGKTNTYNTCYTAGVCLLKARQKI 360
		QY	361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
		Db	361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
		RESULT	10

	Key	Location/Qualifiers
FT	Peptide	1..15 /note= "HA signal"
FT	Region	18..243 /note= "hiG1 Fc aa248-473 of gi2765420"
FT	Peptide	246..259 /note= "linker peptide"
FT	Region	266..410 /note= "hEDA2 aa245-389"
XX		
PX	DE10205583-A1.	
XX		
PD	21-AUG-2003.	
PX	11-FEB-2002; 2002DE-01005583.	
PR	10-FEB-2002; 2002DE-01005368.	
XX	(APOX-) APOXIS SA.	
PI	Gaide O, Schneider P, Tschopp J;	
XX	WPI: 2003-647319/62.	
N-	PSDB; ADP75165.	
New	recombinant fusion protein, useful for treating deficiency of tumor necrosis factor ligand, particularly ectodermal dysplasia, comprises the ligand and immunoglobulin fragment.	
Claim	8; Fig B; 2pp; German.	
This	invention describes a novel recombinant fusion protein which comprises an immunoglobulin Fc segment, the extracellular region of a tumour necrosis factor (TNF) and optionally a transition region, including a linker. The TNF ligands used in the method include FasL, TNF-alpha or TNF-gamma, TRAIL or most preferred, EDAl or 2. The Fc fragment contains the hinge region and domains CH2 and CH3 of the Fc segment of an IgG, particularly human. The linker optionally includes a protease cleavage site. The N-terminus of construct may have a signal sequence, secretory signal sequence and/or tag (e.g. Flag or His). The products of the invention have dermatological and depilatory activity. The constructs or other therapeutic agents, are administered parentally (especially intravenously or intra-arterially) during pregnancy, especially as early as possible, e.g. for humans, in the first 3 months, every second day for at least 14 days. Nucleic acid encoding the construct, related vectors and transfected host cells, are useful in gene/cellular therapy of genetic disorders associated with defective expression of TNF ligand, particularly ectodermal dysplasia (specifically the X-linked hypohydrotic form or X-linked immune deficiency with hyper Igm) and also alopecia, hirsutism and inadequate functioning of the sweat and sebaceous glands. The fusion protein construct is soluble but still retains ability to mimic the physiological effects of TNF ligand. This sequence represents the fusion construct FC-EDAl.	
Query Match	37.1%; Score 779; DB 7; Length 412;	
Best Local Similarity	87.6%; Pred. No. 4e-52;	
Matches	155; Conservative 1; Mismatches 15; Indels 6; Gaps 2;	
QY	215 GPFGPPGGPGPGGLQGSGAAGKATRENQPAAVHLLQQGSIAIQKNDSLGGVLNDW 274 : Dk	
Dd	242 GKRSPOPKPKPKP---EPEGLS Q ---VDENQPAVVHLQGSIAIQKNDSLGGVLNDW 295 :	
Qy	275 SRITMNPVKFKLPKRSEGLELVLDGTFFIYSQVEVIYNFTDFASYEVVDDKEFLQCTR 334 : Db	
Dd	296 SRITMNPVKFKLPKRSEGLELVLDGTFFIYSQ--YYINFTDFASYEVVDDKEFLQCTR 355 :	
Qy	335 SIETGNTNVTCTTAGVCCLLKAROKIAVKWHADISINMSKHHTTFFGAIRLGEPAPAS 391 : Db	
Dd	356 SIETGNTNVTCTTAGVCCLLKAROKIAVKWHADISINMSKHHTTFFGAIRLGEPAPAS 412 :	
RESULT 13		
ID	ADP75166	
XZ	ADP75166 standard; protein; 410 AA.	
AC	ADP75166;	
XX		
DT	12-AUG-2004 (first entry)	
DX	Fusion construct psl236 Fc:EDA2 protein.	
XX		
KW	fusion protein; immunoglobulin; Fc region; TNF; tumour necrosis factor; FasL; TNF-alpha; TNF-gamma; TRAIL; EDA1; EDA2; dermatological; depilatory; gene therapy; cellular therapy; ectodermal dysplasia; X-linked immune deficiency; hyper Igm; alopecia; hirsutism; sweat gland; sebaceous gland. Homo sapiens. Synthetic.	

RESULT 14
AAW44764
ID AAW44764 standard; protein; 135 AA.
XX AC AAW44764;
XX DT 01-JUN-1998 (first entry)
XX DE Human anhidrotic ectodermal dysplasia protein.
XX KW Human; anhidrotic ectodermal dysplasia; X chromosome; genetic linkage;
XX KW translocation; CpG island; foetal development; hair; sweat gland; tooth.
XX OS Homo sapiens.
XX PN US5700926-A.
XX PD 23-DEC-1997.
XX PF 22-JUL-1996; 96US-00684672.
XX PR 27-APR-1993; 93US-00052997.
XX PA (UNIW) UNIV WASHINGTON.
XX PI De La Chapelle A, Srivastava AK, Kere J, Schlessinger D;
XX WIPI; 1998-062436/06.
XX DR N-PSDB; AAV05851.
XX PT Human anhidrotic ectodermal dysplasia gene - useful for research into
XX PT hair growth.
XX PS Disclosure; Fig 7; 37pp; English.
XX CC This is the amino acid sequence of the protein encoded by the human
XX CC anhidrotic ectodermal dysplasia (EDA) gene. The gene has been mapped to
XX CC the region Xq12-q13 by genetic linkage analysis and has been shown to
XX CC contain a 200 kb intron inserted in the 3' end of the coding sequence.
XX CC Deficiencies in the gene are observed by translocations with a breakpoint
XX CC in the transcribed CpG island 3 at the Xq12-q13 locus. The EDA gene can
XX CC be used to study the dynamics of EDA gene expression during foetal
XX CC development, and processes affecting normal hair growth in adults. The
XX CC EDA gene can also be used to study hair, sweat gland and tooth formation
XX CC and growth, and ectodermal dysplasias
XX SQ Sequence 135 AA;
Query Match 33.4%; Score 700; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGPFGLSLALHLITLCC 60
DB 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGPFGLSLALHLITLCC 60
QY 61 YLELSELRRERGAESRLGSGTGTSGTSLSSGLDPSDPTSHLQSPKQOPLPGGE 120
DB 61 YLELSELRRERGAESRLGSGTGTSGTSLSSGLDPSDPTSHLQSPKQOPLPGGE 120
QY 121 AALHSDSDQGHQ 132
DB 121 AALHSDSDQGHQ 132
RESULT 15
AAR95115
ID AAR95115 standard; protein; 357 AA.
XX AC AAR95115;
XX DT 25-MAR-2003 (revised)
XX DT 03-FEB-1997 (first entry)

XX Intervening sequence contg. collagen like protein (CLP)-CB.
XX KW Polymer; repeat unit; natural collagen; intervening oligopeptide; fibre;
XX KW film; membrane; emulsion; coating; collagen like protein; fibre;
XX KW specific binding material; catalyst; purification agent; composite;
XX KW laminate; adhesive; cell growth surface; affinity column;
XX KW biological material support; wound dressing; in vivo prosthesis.
XX OS Synthetic.
XX PN US5514581-A.
XX PD 07-MAY-1996.
XX PF 06-NOV-1990; 90US-00609716.
XX PR 04-NOV-1986; 86US-00927258.
XX PR 29-OCT-1987; 87US-00114618.
XX PR 09-NOV-1988; 88US-00269429.
XX PR 07-NOV-1989; 89WO-US005016.
XX PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX PI Cappello J, Ferrari FA;
XX WIPI; 1996-238772/24.
XX DR DNA encoding protein contg. repeated fibroin derived segments - linked by
XX PT oligopeptide with cell adhesion properties useful, e.g. in wound
XX PT dressings.
XX PS Example 3; Col 101-104; 71pp; English.
XX CC Novel DNA sequence encodes a polymer comprising segments of repeating
XX CC units of 3-9 amino acids from natural collagen, i.e. the present
XX CC intervening sequence contg. collagen like protein (CLP)-CB, able to
XX CC assemble into aligned structures formable into articles. The polymer
XX CC comprises at least 2 segments joined by an unaligned intervening
XX CC oligopeptide, other than the repeating unit. The polymer can be used to
XX CC make fibres, films, membranes, emulsions, coatings, etc., useful as, e.g.
XX CC specific binding materials, catalysts, purificn. agents, composites,
XX CC laminates, adhesives, cell growth surfaces, affinity columns and supports
XX CC for biological materials. Typical applications include wound dressings,
XX CC and in vivo prostheses. The polymer produces articles with good mechanical
XX CC properties, and the intervening oligopeptide can provide a ligand for
XX CC binding a mol., antibody, etc., or a chemically reactive site for
XX CC coupling to proteins, etc. (Updated on 25-MAR-2003 to correct PF field.)
XX SQ Sequence 357 AA;
Query Match 14.0%; Score 294; DB 2; Length 357;
Best Local Similarity 32.4%; Pred. No. 2.1e-14;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;
QY 71 ERGASRLGGSGTGTSGTSLSSGLDPSDPTSHLQSPKQOPLFGAALHSDSQ 128
DB 71 DRGDAGPKGADSGPAGPVGSPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQWALLNFFPDEKPYSEESRVRNRKRSNEGADGPVKNKKKAGKAGPGPGNGPP 188
DB 131 D-----RGDAGPKGADSGPAGPVGSPGPGPGPGPGPGPGPGPGPP 165
QY 189 GPPG 212
DB 166 GPPG 225
QY 213 -VMGPPG 244
DB 226 GPPG 258
Search completed: June 2, 2006, 17:24:59

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:25:22 ; Search time 41 Seconds
(without alignments)
917.579 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRR.....NMSKHTTFFCAIRLGEAPAS 391

Scoring table: GAPUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280.5	13.4	1838	1 CGHU1V	collagen alpha 1(V
2	277	13.2	1464	2 S59856	collagen alpha 1(I
3	275.5	13.1	1433	2 A46053	bullous pemphigoid
4	275	13.1	1049	1 CGB078	collagen alpha 1(I
5	273.5	13.0	1843	2 S18803	collagen alpha 1(V
6	273	13.0	1532	2 A61262	collagen alpha 1(X
7	268.5	12.8	742	2 JC7595	scavenger receptor
8	267.5	12.8	316	2 T20497	hypothetical prote
9	267	12.7	886	2 I50694	collagen alpha 1(I
10	267	12.7	1315	2 A56101	collagen alpha 1(X
11	267	12.7	1774	2 B56101	collagen alpha 1(X
12	266	12.7	779	1 CGB01S	collagen alpha 1(I
13	265.5	12.7	1464	1 CGHU1S	collagen alpha 1(I
14	265	12.6	1763	2 S16366	collagen alpha 2(I
15	264	12.6	1690	1 CGBU1B	collagen alpha 4(I
16	263.5	12.6	1146	2 A38587	collagen, cornea-s
17	263.5	12.6	1419	2 A41182	collagen alpha 1(S
18	263.5	12.6	1487	2 B41182	collagen alpha 1(I
19	263	12.5	1366	1 CGHU2S	collagen alpha 2(I
20	260.5	12.4	675	2 S20819	collagen alpha 3(I
21	259.5	12.4	302	2 T15936	hypothetical prote
22	259.5	12.4	674	2 S23297	collagen alpha 1(X
23	258.5	12.3	1373	1 A43291	collagen alpha 2(I
24	258	12.3	674	2 S13301	collagen alpha 1(X
25	258	12.3	680	1 CGHU1D	collagen alpha 1(X
26	258	12.3	1142	2 JX0369	collagen alpha 1(X
27	257.5	12.3	312	2 T25048	hypothetical prote
28	257.5	12.3	1453	2 S21626	collagen alpha 1(I
29	256.5	12.2	688	2 A53330	collagen alpha 2(I

RESULT 1

CGHU1V

collagen alpha 1(V) chain precursor - human
N;Alternate names: procollagen alpha 1(V) chain

C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence revision 03-Oct-1995 #text change 09-Jul-2004

C;Accession: S18802; S16024; A61142; S11303; S03978; S43642; S58665

R;GreenSpan, D.S.; Cheng, W.; Hoffman, G.G.

J. Biol. Chem. 266, 24727-24733, 1991

A;Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution of e

A;Reference number: S18802; MUID:92105142; PMID:1722213

A;Accession: S18802

A;Molecule type: mRNA

A;Residues: 1-1838 <GRE>

A;Cross-references: UNIPARC:Q15094; UNIPARC:UPI0000065982; GB:M76729; NID:q189519; PIDN:R

R;Takahara, K.; Sato, Y.; Okazawa, K.; Okamoto, N.; Noda, A.; Yaoi, Y.; Kato, I.

J. Biol. Chem. 266, 13124-13129, 1991

A;Title: Complete primary structure of human collagen alpha-1(V) chain.

A;Reference number: S16024; MUID:91302336; PMID:2071595

A;Accession: S16024

A;Molecule type: mRNA

A;Residues: 1-81, QL', 84-389, 'A', 391-676, 'K', 678-1294, 'PS', 1297, 'RS', 1300-1553, 'R', 1555-J

A;Cross-references: UNIPARC:UPI0000126D1D; GB:D90279; NID:G219509; PIDN:BAAL4323.1; PID:R

R;Yaoi, Y.; Hashimoto, K.; Takahara, K.; Kato, I.

Exp. Cell Res. 194, 180-185, 1991

A;Title: Insulin binds to type V collagen with retention of mitogenic activity.

A;Reference number: A61142; MUID:91224163; PMID:1709100

A;Accession: A61142

A;Molecule type: protein

A;Residues: 823-824, 'X', 826-842 <YAO>

A;Cross-references: UNIPARC:UPI0000173BB7

A;Note: the residue designated 'X' is probably glycosylated hydroxylysine; this cyanogen

R;Yaoi, Y.; Hashimoto, K.; Koitabashi, H.; Takahara, K.; Ito, M.; Kato, I.

Biochim. Biophys. Acta 1035, 139-145, 1990

A;Title: Primary structure of the heparin-binding site of type V collagen.

A;Reference number: S11303; MUID:90366601; PMID:2203476

A;Accession: S11303

A;Molecule type: protein

A;Residues: 823-824, 'X', 826-848, 'I', 850-851, 'P', 853, 'PR', 856-893, 'D', 895-932, 'X', 934-950

A;Cross-references: UNIPARC:UPI0000173BB8

A;Note: the residues designated 'X' are probably glycosylated hydroxylysine; this sequen

R;Seyer, J.M.; Kang, A.H.

Arch. Biochem. Biophys. 271, 120-129, 1989

A;Title: Covalent structure of collagen: amino acid sequence of three cyanogen bromide-d

A;Reference number: S03978; MUID:89227189; PMID:2496661

A;Accession: S03978

A;Molecule type: protein

A;Residues: 621-640, 'G', 642-649, 'L', 651-662, 'E', 664-667, 'Q', 669-676, 'Q', 678-683, 'P', 685-6

<SE>

A;Cross-references: UNIPARC:UPI0000173BB9

A;Note: there are a number of inconsistencies between the sequences in figures 6 and 7; t

R;Moradi-Ameli, M.; Rousseau, J.C.; Kleman, J.P.; Champliaud, M.F.; Boutillon, M.M.; Ben Eur. J. Biochem. 221, 987-995, 1994
A:Title: Diversity in the processing events at the N-terminus of type-V collagen.
A:Reference number: S43642; MUID:94237164; PMID:8181482
A:Accession: S43642
A:Molecule type: protein
A:Residues: 565-576;756-758, 'X', 760-763, 'X', 765-772;1012-1029;1219-1232;1465-1474, 'X', 1474-1482
A:Cross-references: UNIPARC:UPI0000173BBA; UNIPARC:UPI0000173BBB; UNIPARC:UPI0000173BBC;
R;Fessler, L.I.; Brosh, S.; Chapin, S.; Fessler, J.H.
J. Biol. Chem. 261, 5034-5040, 1986
A:Title: Tyrosine sulfation in precursors of collagen V.
A:Reference number: A56977; MUID:86168226; PMID:3082875
A:Contents: annotation; identification of tyrosine sulfate in the amino-terminal propeptide of collagen V.
R;Lee, S.; Greenspan, D.S.
Biochem. J. 310, 15-22, 1995
A:Title: Transcriptional promoter of the human alpha-1(V) collagen gene (COL5A1).
A:Reference number: S58665; MUID:95374437; PMID:7646438
A:Accession: S58665
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-36 <LRE>
A:Cross-references: UNIPARC:UPI00007363D; GB:L38808; NID:G1020325; PIDN:AAA79853.1; PIDN:AAA79853.2
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit are 5-hydroxylated and subsequently O-glycosylated.
C:Comment: A long form of the mature protein containing part of the amino-terminal propeptide the heterotrimers are probably processed to the long form.
C:Genetics:
A:Gene: GDB:COL5A1
A:Cross-references: GDB:131457; OMIM:120215
A:Map position: 9q34.2-9q34.3
C:Complex: type V collagen may be a homotrimer of alpha 1(V) chains, a heterotrimer of alpha 2(V) chain and one alpha 3(V) chain, initially linked by disulfide bonds among the chains, is formed with desmosine cross-links made from lysine and allysine residues
C:Function:
A:Description: structural component of extracellular fibrous polymer associated with cell attachment
A:Note: may play a role in controlling the lateral growth of collagen I fibrils
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
F:1-37/Domain: signal sequence #status predicted <SIG>
F:36-261/Domain: PAMP-like #status predicted <PAMP>
F:38-541/Domain: amino-terminal propeptide #status predicted <PRO>
F:542-1605/Product: collagen alpha 1(V) chain, short form #status predicted <MAT>
F:542-558/Region: amino-terminal nonhelical telopeptide
F:559-1572/Region: helical
F:645-647/Region: cell attachment (R-G-D) motif
F:663-665/Region: cell attachment (R-G-D) motif
F:897-929/Region: heparin binding
F:1573-1605/Region: carboxyl-terminal nonhelical telopeptide
F:1606-1838/Domain: carboxyl-terminal propeptide #status predicted <CPR>
F:1615-1837/Domain: fibrillar collagen carboxylic acid (Gln) (in mature form) #status predicted <FCC>
F:38/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:62-244,183-237/Disulfide bonds: #status predicted
F:159,176,385,1672,1741/Binding site: carboxylate (Asn) (covalent) #status predicted
F:234,236,240,262,263,273,274,275,277,280,338,340,346,347,352,357,416,417,420,421/Binding site: allysine (Lys) #status predicted
F:535/Modified site: allysine (Lys) #status predicted
F:541-542/Cleavage site: Ala-Gln (procollagen N-endorpeptidase) #status predicted
F:542/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:570,576,622,639,648,654,657,675,678,696,693,696,705,717,720,726,732,741,750,753,756,760/Modified site: 4-hydroxyproline (Pro) #status experimental
F:627,642,687,708,744,774,795,804,807,810,819,825,846,864,882,897/Modified site: 5-hydroxyproline (Pro) #status experimental
F:627,642,687,774,795,804,807,810,819,825,846,864,882,897,1482/Binding site: carboxylate (Asn) (covalent) #status experimental
F:708,744/Binding site: carboxylate (Lys) (covalent) #status predicted
F:1482/Modified site: 5-hydroxylysine (Lys) #status predicted
F:1605-1606/cleavage site: Ala-Asp (procollagen C-endorpeptidase) #status predicted
F:1639,1645,1662,1671/Disulfide bonds: interchain #status predicted
F:1680-1835,1746-1789/Disulfide bonds: #status predicted

Db 857 PLGPPGEEKGLGVPLGPGVGRGPKGS-----IGPFGFP----- 891
QY 72 RGAESRLGGSGTSGTSLSSGLDPSDPSITSHLGQSPSPKQPLPGEAALHSDSQDGH 131
Db 892 -GANEKGGRTG-----GKPGRGGRGPTG----- 916
QY 132 QWALLNFFPDDBKPYSEESRRVRNRKSKSNEGADGPGVKNKKKGKAGPP---GNNGPP 188
Db 917 -----PRGERPGRTGKPGKSGSGDGP-----AGPPGERGPNGPQ 954
QY 189 GPPGPPGPGPGIGIPGIPG-----TTVMGPPGPPGP----- 222
Db 955 GTGTFGPGKPGPGPKDGLPGHGPQGRGTGQKGTGPGVGVGPGQGTGTGTGPMGR 1014
QY 223 --PGQGPGLGPGSGAADKATRENQPAVVHLOGQ 256
Db 1015 GHPPGPPPGEGGLGGLAGKGTGK-GDQPGAGLPGK 1049

RESULT 2
S59856
collagen alpha 1(III) chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C:Accession: S59856; S62120; S16373
R:Toman, P.D.; de Crombrughe, B.
Gene 147, 161-168, 1994
A:Title: The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA sequence.
A:Reference number: S59856; MUID:95011609; PMID:7926795
A:Accession: S59856
A:Molecule type: DNA
A:Residues: 1-1464 <TO>
A:Cross-references: UNIPROT:P08121; UNIPARC:UPI0000177386; EMBL:X52046
R:Toman, D.
Submitted to the EMBL Data Library, November 1994
A:Reference number: S62120
A:Accession: S62120
A:Molecule type: DNA
A:Residues: 1-866, 'G', 868-1464 <TO>
A:Cross-references: UNIPARC:UPI0000295D6; EMBL:X52046; NID:G575321; PIDN:CAA36279.1; PIDN:CAA36279.2
R:Metzger, M.; Roman, D.; de Crombrughe, B.; Vuorio, E.
Biochim. Biophys. Acta 1089, 241-243, 1991
A:Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
A:Reference number: S16176; MUID:91274355; PMID:2054384
A:Accession: S16373
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1442-1464 <MET>
A:Cross-references: UNIPARC:UPI000016CCAB; EMBL:X57983; NID:G50476; PIDN:CAA41048.1; PIDN:CAA41048.2
C:Genetics:
A:Introns: 29/1; 95/3; 112/3; 150/3; 175/3; 193/3; 211/3; 229/3; 247/3; 265/3; 283/3; 298/3; 316/3; 334/3; 352/3; 370/3; 388/3; 406/3; 424/3; 442/3; 460/3; 478/3; 496/3; 514/3; 532/3; 550/3; 568/3; 586/3; 604/3; 622/3; 640/3; 658/3; 676/3; 694/3; 712/3; 730/3; 748/3; 766/3; 784/3; 802/3; 820/3; 838/3; 856/3; 874/3; 892/3; 910/3; 928/3; 946/3; 964/3; 982/3; 1000/3; 1018/3; 1036/3; 1054/3; 1072/3; 1090/3; 1108/3; 1126/3; 1144/3; 1162/3; 1180/3; 1198/3; 1216/3; 1234/3; 1252/3; 1270/3; 1288/3; 1306/3; 1324/3; 1342/3; 1360/3; 1378/3; 1396/3; 1414/3; 1432/3; 1450/3; 1468/3; 1486/3; 1504/3; 1522/3; 1540/3; 1558/3; 1576/3; 1594/3; 1612/3; 1630/3; 1648/3; 1666/3; 1684/3; 1702/3; 1720/3; 1738/3; 1756/3; 1774/3; 1792/3; 1810/3; 1828/3; 1846/3; 1864/3; 1882/3; 1900/3; 1918/3; 1936/3; 1954/3; 1972/3; 1990/3; 2008/3; 2026/3; 2044/3; 2062/3; 2080/3; 2098/3; 2116/3; 2134/3; 2152/3; 2170/3; 2188/3; 2206/3; 2224/3; 2242/3; 2260/3; 2278/3; 2296/3; 2314/3; 2332/3; 2350/3; 2368/3; 2386/3; 2404/3; 2422/3; 2440/3; 2458/3; 2476/3; 2494/3; 2512/3; 2530/3; 2548/3; 2566/3; 2584/3; 2602/3; 2620/3; 2638/3; 2656/3; 2674/3; 2692/3; 2710/3; 2728/3; 2746/3; 2764/3; 2782/3; 2800/3; 2818/3; 2836/3; 2854/3; 2872/3; 2890/3; 2908/3; 2926/3; 2944/3; 2962/3; 2980/3; 2998/3; 3016/3; 3034/3; 3052/3; 3070/3; 3088/3; 3106/3; 3124/3; 3142/3; 3160/3; 3178/3; 3196/3; 3214/3; 3232/3; 3250/3; 3268/3; 3286/3; 3304/3; 3322/3; 3340/3; 3358/3; 3376/3; 3394/3; 3412/3; 3430/3; 3448/3; 3466/3; 3484/3; 3502/3; 3520/3; 3538/3; 3556/3; 3574/3; 3592/3; 3610/3; 3628/3; 3646/3; 3664/3; 3682/3; 3700/3; 3718/3; 3736/3; 3754/3; 3772/3; 3790/3; 3808/3; 3826/3; 3844/3; 3862/3; 3880/3; 3898/3; 3916/3; 3934/3; 3952/3; 3970/3; 3988/3; 4006/3; 4024/3; 4042/3; 4060/3; 4078/3; 4096/3; 4114/3; 4132/3; 4150/3; 4168/3; 4186/3; 4204/3; 4222/3; 4240/3; 4258/3; 4276/3; 4294/3; 4312/3; 4330/3; 4348/3; 4366/3; 4384/3; 4402/3; 4420/3; 4438/3; 4456/3; 4474/3; 4492/3; 4510/3; 4528/3; 4546/3; 4564/3; 4582/3; 4600/3; 4618/3; 4636/3; 4654/3; 4672/3; 4690/3; 4708/3; 4726/3; 4744/3; 4762/3; 4780/3; 4798/3; 4816/3; 4834/3; 4852/3; 4870/3; 4888/3; 4906/3; 4924/3; 4942/3; 4960/3; 4978/3; 4996/3; 5014/3; 5032/3; 5050/3; 5068/3; 5086/3; 5104/3; 5122/3; 5140/3; 5158/3; 5176/3; 5194/3; 5212/3; 5230/3; 5248/3; 5266/3; 5284/3; 5302/3; 5320/3; 5338/3; 5356/3; 5374/3; 5392/3; 5410/3; 5428/3; 5446/3; 5464/3; 5482/3; 5500/3; 5518/3; 5536/3; 5554/3; 5572/3; 5590/3; 5608/3; 5626/3; 5644/3; 5662/3; 5680/3; 5698/3; 5716/3; 5734/3; 5752/3; 5770/3; 5788/3; 5806/3; 5824/3; 5842/3; 5860/3; 5878/3; 5896/3; 5914/3; 5932/3; 5950/3; 5968/3; 5986/3; 6004/3; 6022/3; 6040/3; 6058/3; 6076/3; 6094/3; 6112/3; 6130/3; 6148/3; 6166/3; 6184/3; 6202/3; 6220/3; 6238/3; 6256/3; 6274/3; 6292/3; 6310/3; 6328/3; 6346/3; 6364/3; 6382/3; 6400/3; 6418/3; 6436/3; 6454/3; 6472/3; 6490/3; 6508/3; 6526/3; 6544/3; 6562/3; 6580/3; 6598/3; 6616/3; 6634/3; 6652/3; 6670/3; 6688/3; 6706/3; 6724/3; 6742/3; 6760/3; 6778/3; 6796/3; 6814/3; 6832/3; 6850/3; 6868/3; 6886/3; 6904/3; 6922/3; 6940/3; 6958/3; 6976/3; 6994/3; 7012/3; 7030/3; 7048/3; 7066/3; 7084/3; 7102/3; 7120/3; 7138/3; 7156/3; 7174/3; 7192/3; 7210/3; 7228/3; 7246/3; 7264/3; 7282/3; 7300/3; 7318/3; 7336/3; 7354/3; 7372/3; 7390/3; 7408/3; 7426/3; 7444/3; 7462/3; 7480/3; 7498/3; 7516/3; 7534/3; 7552/3; 7570/3; 7588/3; 7606/3; 7624/3; 7642/3; 7660/3; 7678/3; 7696/3; 7714/3; 7732/3; 7750/3; 7768/3; 7786/3; 7804/3; 7822/3; 7840/3; 7858/3; 7876/3; 7894/3; 7912/3; 7930/3; 7948/3; 7966/3; 7984/3; 8002/3; 8020/3; 8038/3; 8056/3; 8074/3; 8092/3; 8110/3; 8128/3; 8146/3; 8164/3; 8182/3; 8200/3; 8218/3; 8236/3; 8254/3; 8272/3; 8290/3; 8308/3; 8326/3; 8344/3; 8362/3; 8380/3; 8398/3; 8416/3; 8434/3; 8452/3; 8470/3; 8488/3; 8506/3; 8524/3; 8542/3; 8560/3; 8578/3; 8596/3; 8614/3; 8632/3; 8650/3; 8668/3; 8686/3; 8704/3; 8722/3; 8740/3; 8758/3; 8776/3; 8794/3; 8812/3; 8830/3; 8848/3; 8866/3; 8884/3; 8902/3; 8920/3; 8938/3; 8956/3; 8974/3; 8992/3; 9010/3; 9028/3; 9046/3; 9064/3; 9082/3; 9100/3; 9118/3; 9136/3; 9154/3; 9172/3; 9190/3; 9208/3; 9226/3; 9244/3; 9262/3; 9280/3; 9298/3; 9316/3; 9334/3; 9352/3; 9370/3; 9388/3; 9406/3; 9424/3; 9442/3; 9460/3; 9478/3; 9496/3; 9514/3; 9532/3; 9550/3; 9568/3; 9586/3; 9604/3; 9622/3; 9640/3; 9658/3; 9676/3; 9694/3; 9712/3; 9730/3; 9748/3; 9766/3; 9784/3; 9802/3; 9820/3; 9838/3; 9856/3; 9874/3; 9892/3; 9910/3; 9928/3; 9946/3; 9964/3; 9982/3; 10000/3; 10018/3; 10036/3; 10054/3; 10072/3; 10090/3; 10108/3; 10126/3; 10144/3; 10162/3; 10180/3; 10198/3; 10216/3; 10234/3; 10252/3; 10270/3; 10288/3; 10306/3; 10324/3; 10342/3; 10360/3; 10378/3; 10396/3; 10414/3; 10432/3; 10450/3; 10468/3; 10486/3; 10504/3; 10522/3; 10540/3; 10558/3; 10576/3; 10594/3; 10612/3; 10630/3; 10648/3; 10666/3; 10684/3; 10702/3; 10720/3; 10738/3; 10756/3; 10774/3; 10792/3; 10810/3; 10828/3; 10846/3; 10864/3; 10882/3; 10900/3; 10918/3; 10936/3; 10954/3; 10972/3; 10990/3; 11008/3; 11026/3; 11044/3; 11062/3; 11080/3; 11098/3; 11116/3; 11134/3; 11152/3; 11170/3; 11188/3; 11206/3; 11224/3; 11242/3; 11260/3; 11278/3; 11296/3; 11314/3; 11332/3; 11350/3; 11368/3; 11386/3; 11404/3; 11422/3; 11440/3; 11458/3; 11476/3; 11494/3; 11512/3; 11530/3; 11548/3; 11566/3; 11584/3; 11602/3; 11620/3; 11638/3; 11656/3; 11674/3; 11692/3; 11710/3; 11728/3; 11746/3; 11764/3; 11782/3; 11800/3; 11818/3; 11836/3; 11854/3; 11872/3; 11890/3; 11908/3; 11926/3; 11944/3; 11962/3; 11980/3; 11998/3; 12016/3; 12034/3; 12052/3; 12070/3; 12088/3; 12106/3; 12124/3; 12142/3; 12160/3; 12178/3; 12196/3; 12214/3; 12232/3; 12250/3; 12268/3; 12286/3; 12304/3; 12322/3; 12340/3; 12358/3; 12376/3; 12394/3; 12412/3; 12430/3; 12448/3; 12466/3; 12484/3; 12502/3; 12520/3; 12538/3; 12556/3; 12574/3; 12592/3; 12610/3; 12628/3; 12646/3; 12664/3; 12682/3; 12700/3; 12718/3; 12736/3; 12754/3; 12772/3; 12790/3; 12808/3; 12826/3; 12844/3; 12862/3; 12880/3; 12898/3; 12916/3; 12934/3; 12952/3; 12970/3; 12988/3; 13006/3; 13024/3; 13042/3; 13060/3; 13078/3; 13096/3; 13114/3; 13132/3; 13150/3; 13168/3; 13186/3; 13204/3; 13222/3; 13240/3; 13258/3; 13276/3; 13294/3; 13312/3; 13330/3; 13348/3; 13366/3; 13384/3; 13402/3; 13420/3; 13438/3; 13456/3; 13474/3; 13492/3; 13510/3; 13528/3; 13546/3; 13564/3; 13582/3; 13600/3; 13618/3; 13636/3; 13654/3; 13672/3; 13690/3; 13708/3; 13726/3; 13744/3; 13762/3; 13780/3; 13798/3; 13816/3; 13834/3; 13852/3; 13870/3; 13888/3; 13906/3; 13924/3; 13942/3; 13960/3; 13978/3; 13996/3; 14014/3; 14032/3; 14050/3; 14068/3; 14086/3; 14104/3; 14122/3; 14140/3; 14158/3; 14176/3; 14194/3; 14212/3; 14230/3; 14248/3; 14266/3; 14284/3; 14302/3; 14320/3; 14338/3; 14356/3; 14374/3; 14392/3; 14410/3; 14428/3; 14446/3; 14464/3; 14482/3; 14500/3; 14518/3; 14536/3; 14554/3; 14572/3; 14590/3; 14608/3; 14626/3; 14644/3; 14662/3; 14680/3; 14698/3; 14716/3; 14734/3; 14752/3; 14770/3; 14788/3; 14806/3; 14824/3; 14842/3; 14860/3; 14878/3; 14896/3; 14914/3; 14932/3; 14950/3; 14968/3; 14986/3; 15004/3; 15022/3; 15040/3; 15058/3; 15076/3; 15094/3; 15112/3; 15130/3; 15148/3; 15166/3; 15184/3; 15202/3; 15220/3; 15238/3; 15256/3; 15274/3; 15292/3; 15310/3; 15328/3; 15346/3; 15364/3; 15382/3; 15400/3; 15418/3; 15436/3; 15454/3; 15472/3; 15490/3; 15508/3; 15526/3; 15544/3; 15562/3; 15580/3; 15598/3; 15616/3; 15634/3; 15652/3; 15670/3; 15688/3; 15706/3; 15724/3; 15742/3; 15760/3; 15778/3; 15796/3; 15814/3; 15832/3; 15850/3; 15868/3; 15886/3; 15904/3; 15922/3; 15940/3; 15958/3; 15976/3; 15994/3; 16012/3; 16030/3; 16048/3; 16066/3; 16084/3; 16102/3; 16120/3; 16138/3; 16156/3; 16174/3; 16192/3; 16210/3; 16228/3; 16246/3; 16264/3; 16282/3; 16300/3; 16318/3; 16336/3; 16354/3; 16372/3; 16390/3; 16408/3; 16426/3; 16444/3; 16462/3; 16480/3; 16498/3; 16516/3; 16534/3; 16552/3; 16570/3; 16588/3; 16606/3; 16624/3; 16642/3; 16660/3; 16678/3; 16696/3; 16714/3; 16732/3; 16750/3; 16768/3; 16786/3; 16804/3; 16822/3; 16840/3; 16858/3; 16876/3; 16894/3; 16912/3; 16930/3; 16948/3; 16966/3; 16984/3; 17002/3; 17020/3; 17038/3; 17056/3; 17074/3; 17092/3; 17110/3; 17128/3; 17146/3; 17164/3; 17182/3; 17200/3; 17218/3; 17236/3; 17254/3; 17272/3; 17290/3; 17308/3; 17326/3; 17344/3; 17362/3; 17380/3; 17398/3; 17416/3; 17434/3; 17452/3; 17470/3; 17488/3; 17506/3; 17524/3; 17542/3; 17560/3; 17578/3; 17596/3; 17614/3; 17632/3; 17650/3; 17668/3; 17686/3; 17704/3; 17722/3; 17740/3; 17758/3; 17776/3; 17794/3; 17812/3; 17830/3; 17848/3; 17866/3; 17884/3; 17902/3; 17920/3; 17938/3; 17956/3; 17974/3; 17992/3; 18010/3; 18028/3; 18046/3; 18064/3; 18082/3; 18100/3; 18118/3; 18136/3; 18154/3; 18172/3; 18190/3; 18208/3; 18226/3; 18244/3; 18262/3; 18280/3; 18298/3; 18316/3; 18334/3; 18352/3; 18370/3; 18388/3; 18406/3; 18424/3; 18442/3; 18460/3; 18478/3; 18496/3; 18514/3; 18532/3; 18550/3; 18568/3; 18586/3; 18604/3; 18622/3; 18640/3; 18658/3; 18676/3; 18694/3; 18712/3; 18730/3; 18748/3; 18766/3; 18784/3; 18802/3; 18820/3; 18838/3; 18856/3; 18874/3; 18892/3; 18910/3; 18928/3; 18946/3; 18964/3; 18982/3; 19000/3; 19018/3; 19036/3; 19054/3; 19072/3; 19090/3; 19108/3; 19126/3; 19144/3; 19162/3; 19180/3; 19198/3; 19216/3; 19234/3; 19252/3; 19270/3; 19288/3; 19306/3; 19324/3; 19342/3; 19360/3; 19378/3; 19396/3; 19414/3; 19432/3;

Eur. J. Biochem. 30, 169-183, 1972

A;Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues

A;Reference number: A91201; MUID:73042276; PMID:4343808

A;Accession: A91201

A:Molecule type: protein

A;Residues: 563-675 <WEN>

A;Cross-references: UNIPARC:UPI0000173B55

A;Experimental source: skin

R;Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.

Eur. J. Biochem. 30, 163-168, 1972

A;Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C;

A;Reference number: A91200; MUID:73042275; PMID:4343807

A;Accession: A91200

A:Molecule type: protein

A;Residues: 676-758 <FI4>

A;Cross-references: UNIPARC:UPI0000173B56

A;Experimental source: skin

A;Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in position 726.

R;Rautenberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.

FEBS Lett. 21, 75-79, 1972

A;Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of type I collagen.

A;Reference number: A43048

A;Accession: A43048

A:Molecule type: protein

A;Residues: 759-779 <RA2>

A;Cross-references: UNIPARC:UPI0000173B57

A;Experimental source: skin

C;Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxylated.

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C;Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer; C;Comment: The order of the eight CNBr peptides in the alpha 1(I) chain of bovine skin collagen is: 1, 149, 268, and 217 residues.

C;Comment: The complete chain contains 1052 residues.

C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology; C;Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglutamic acid; trimer; F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 12.7%; Score 266; DB 1; Length 779;

Best Local Similarity 30.0%; Pred. No. 6.1e-10;

Matches 78; Conservative 16; Mismatches 80; Indels 86; Gaps 9;

QY 13 PAAAPRRSGQ---CGCGAPARAG---GNSCLFLGFFGLSLALHLTLCCYLEL 64

DB 339 PAGAPGDKGAGSGPGATGAPGRGEPGPGPA-----GFG----- 377

QY 65 RSELRERGAESRLGSGTGTCSTLSLGLDPSITSLGQSPKQQLPGEAALH 124

DB 378 -----PPGADGQPGAKGFGDAKGDAGPPGAGP---AGPPGPIGNVAGPG--- 423

QY 125 SDSQDGHQALLNFFPFDEKPYSEESRRVRNRKSKSNEGADGVYKNNKKKKAGPPGP 184

DB 424 -----KGAGSAGPPGATGFPGAAGRVGPPGP 450

QY 185 NGPPGPPGPPG-----QGPPGIPGIPGITTVMGPPGPPGPPGPPGLOGPSGAA 238

DB 451 SGNAGPPGPPGAGKEGSKGRGTGPAGRGE---VGPPGPPGAGKAGPADGAPAG 508

QY 239 DKAGTR--ENQPAVVHQQ 256

DB 509 GTTGPQIAGQGVVGLPGQ 528

RESULT 13

CGHUIS

collagen alpha 1(I) chain precursor - human

N;Alternate names: procollagen alpha 1(I) chain

C;Species: Homo sapiens (man)

C;Date: 12-Aug-1981 #sequence revision 04-Oct-1996 #text_change 31-Dec-2004

C;Accession: I60114; S01143; A33335; I55254; A39943; I55237; A35233; S09400; B90567; S11:

S269; A29439; I53466; A02852; I37247

R;D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.

Gene 67, 105-115, 1998

A;Title: Complete nucleotide sequence of the region encompassing the first twenty-five exons of the human alpha 1(I) collagen gene.

A;Reference number: I60114; MUID:88329734; PMID:2843432

A:Accession: I60114
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-369, 'L', 371-589 <DAL>
A:Cross-references: UNIPROT:P02452; UNIPROT:Q14992; UNIPROT:Q16053; UNIPROT:Q13896; UNIPROT:Q13897; Kuivaniemi, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prockop, D.J. Biochem. J. 253, 919-922, 1988
A:Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human procollagen
A:Reference number: S01143; MUID:89025644; PMID:3178743
A:Accession: S01143
A:Molecule type: mRNA
A:Residues: 1-472 <TRO>
A:Cross-references: UNIPARC:UPI000016A6F9; EMBL:X07884; NID:G30015; PIDN:CAA30731.1; PIDN:CAA30731.1; PIDN:CAA30731.1
A:Note: submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.; Nature 310, 337-340, 1984
A:Title: Human proalpha1(I) collagen gene structure reveals evolutionary conservation of the first intron
A:Reference number: A93335; MUID:84270697; PMID:6462220
A:Accession: A93335
A:Molecule type: DNA
A:Residues: 1-58, 'Q', 60-181 <CHU>
A:Cross-references: UNIPARC:UPI0000173B3C; EMBL:X00820; NID:G35657; PIDN:CAA25394.1; PIDN:CAA25394.1; PIDN:CAA25394.1
R:Rosouw, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.; J. Biol. Chem. 262, 15151-15157, 1987
A:Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene enhancer
A:Reference number: I55254; MUID:88033098; PMID:2822714
A:Accession: I55254
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-45 <ROS>
A:Cross-references: UNIPARC:UPI000016A6B2; GB:J02829; NID:G180387; PIDN:AAA51993.1; PIDN:AAA51993.1; PIDN:AAA51993.1
R:Bornstein, P.; McKay, J.; Morishima, J.K.; Devarayalu, S.; Gelin, R.E. Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987
A:Title: Regulatory elements in the first intron contribute to transcriptional control of the human pro-alpha 1(I) collagen gene
A:Reference number: A39943; MUID:88097389; PMID:3480516
A:Accession: A39943
A:Molecule type: DNA
A:Residues: 1-34 <BOR>
A:Cross-references: UNIPARC:UPI000016A6B1; GB:J03559; NID:G180876; PIDN:AAA52052.1; PIDN:AAA52052.1; PIDN:AAA52052.1
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F. J. Biol. Chem. 260, 2315-2320, 1985
A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter and enhancer
A:Reference number: I55237; MUID:85130970; PMID:2857713
A:Accession: I55237
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <CH2>
A:Cross-references: UNIPARC:UPI000016A6B1; GB:M10627; NID:G180383; PIDN:AAA51992.1; PIDN:AAA51992.1; PIDN:AAA51992.1
R:Wirtz, M.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollister, J. Biol. Chem. 265, 6312-6317, 1990
A:Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-terminal domain, type VII.
A:Reference number: A35233; MUID:90202908; PMID:2318855
A:Accession: A35233
A:Molecule type: protein
A:Residues: 33-52 <WIR>
A:Cross-references: UNIPARC:UPI0000173B3D
A:Note: this propeptide fragment remained non-covalently bound to a defective, uncleaved propeptide
R:Wirtz, M.K.; de Wet, W.; Ramirez, F.; de Wet, W.; Cole, W.G.; Chan, D.; Bateman, J.F. EMBO J. 8, 1705-1710, 1989
A:Title: A base substitution in the exon of a collagen gene causes alternative splicing
A:Reference number: S09400; MUID:89356643; PMID:2767050
A:Accession: S09400
A:Molecule type: mRNA
A:Residues: 156-183 <WBI>
A:Cross-references: UNIPARC:UPI0000173B3E
R:Click, E.M.; Bornstein, P. Biochemistry 9, 4699-4706, 1970
A:Title: Isolation and characterization of the cyanogen bromide peptides from the alpha1(I) chain of human procollagen
A:Reference number: A90567; MUID:71038625; PMID:5529814
A:Contents: CNBr0-1, CNBr2, CNBr4, CNBr5
A:Accession: B90567
A:Molecule type: protein

A:Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z', 234, 'Z', 235, 'Z', 236, 'Z', 237, 'Z', 238, 'Z', 239, 'Z', 240, 'Z', 241, 'Z', 242, 'Z', 243, 'Z', 244, 'Z', 245, 'Z', 246, 'Z', 247, 'Z', 248, 'Z', 249, 'Z', 250, 'Z', 251, 'Z', 252, 'Z', 253, 'Z', 254, 'Z', 255, 'Z', 256, 'Z', 257, 'Z', 258, 'Z', 259, 'Z', 260, 'Z', 261, 'Z', 262, 'Z', 263, 'Z', 264, 'Z', 265, 'Z', 266, 'Z', 267, 'Z', 268, 'Z', 269, 'Z', 270, 'Z', 271, 'Z', 272, 'Z', 273, 'Z', 274, 'Z', 275, 'Z', 276, 'Z', 277, 'Z', 278, 'Z', 279, 'Z', 280, 'Z', 281, 'Z', 282, 'Z', 283, 'Z', 284, 'Z', 285, 'Z', 286, 'Z', 287, 'Z', 288, 'Z', 289, 'Z', 290, 'Z', 291, 'Z', 292, 'Z', 293, 'Z', 294, 'Z', 295, 'Z', 296, 'Z', 297, 'Z', 298, 'Z', 299, 'Z', 300, 'Z', 301, 'Z', 302, 'Z', 303, 'Z', 304, 'Z', 305, 'Z', 306, 'Z', 307, 'Z', 308, 'Z', 309, 'Z', 310, 'Z', 311, 'Z', 312, 'Z', 313, 'Z', 314, 'Z', 315, 'Z', 316, 'Z', 317, 'Z', 318, 'Z', 319, 'Z', 320, 'Z', 321, 'Z', 322, 'Z', 323, 'Z', 324, 'Z', 325, 'Z', 326, 'Z', 327, 'Z', 328, 'Z', 329, 'Z', 330, 'Z', 331, 'Z', 332, 'Z', 333, 'Z', 334, 'Z', 335, 'Z', 336, 'Z', 337, 'Z', 338, 'Z', 339, 'Z', 340, 'Z', 341, 'Z', 342, 'Z', 343, 'Z', 344, 'Z', 345, 'Z', 346, 'Z', 347, 'Z', 348, 'Z', 349, 'Z', 350, 'Z', 351, 'Z', 352, 'Z', 353, 'Z', 354, 'Z', 355, 'Z', 356, 'Z', 357, 'Z', 358, 'Z', 359, 'Z', 360, 'Z', 361, 'Z', 362, 'Z', 363, 'Z', 364, 'Z', 365, 'Z', 366, 'Z', 367, 'Z', 368, 'Z', 369, 'Z', 370, 'Z', 371, 'Z', 372, 'Z', 373, 'Z', 374, 'Z', 375, 'Z', 376, 'Z', 377, 'Z', 378, 'Z', 379, 'Z', 380, 'Z', 381, 'Z', 382, 'Z', 383, 'Z', 384, 'Z', 385, 'Z', 386, 'Z', 387, 'Z', 388, 'Z', 389, 'Z', 390, 'Z', 391, 'Z', 392, 'Z', 393, 'Z', 394, 'Z', 395, 'Z', 396, 'Z', 397, 'Z', 398, 'Z', 399, 'Z', 400, 'Z', 401, 'Z', 402, 'Z', 403, 'Z', 404, 'Z', 405, 'Z', 406, 'Z', 407, 'Z', 408, 'Z', 409, 'Z', 410, 'Z', 411, 'Z', 412, 'Z', 413, 'Z', 414, 'Z', 415, 'Z', 416, 'Z', 417, 'Z', 418, 'Z', 419, 'Z', 420, 'Z', 421, 'Z', 422, 'Z', 423, 'Z', 424, 'Z', 425, 'Z', 426, 'Z', 427, 'Z', 428, 'Z', 429, 'Z', 430, 'Z', 431, 'Z', 432, 'Z', 433, 'Z', 434, 'Z', 435, 'Z', 436, 'Z', 437, 'Z', 438, 'Z', 439, 'Z', 440, 'Z', 441, 'Z', 442, 'Z', 443, 'Z', 444, 'Z', 445, 'Z', 446, 'Z', 447, 'Z', 448, 'Z', 449, 'Z', 450, 'Z', 451, 'Z', 452, 'Z', 453, 'Z', 454, 'Z', 455, 'Z', 456, 'Z', 457, 'Z', 458, 'Z', 459, 'Z', 460, 'Z', 461, 'Z', 462, 'Z', 463, 'Z', 464, 'Z', 465, 'Z', 466, 'Z', 467, 'Z', 468, 'Z', 469, 'Z', 470, 'Z', 471, 'Z', 472, 'Z', 473, 'Z', 474, 'Z', 475, 'Z', 476, 'Z', 477, 'Z', 478, 'Z', 479, 'Z', 480, 'Z', 481, 'Z', 482, 'Z', 483, 'Z', 484, 'Z', 485, 'Z', 486, 'Z', 487, 'Z', 488, 'Z', 489, 'Z', 490, 'Z', 491, 'Z', 492, 'Z', 493, 'Z', 494, 'Z', 495, 'Z', 496, 'Z', 497, 'Z', 498, 'Z', 499, 'Z', 500, 'Z', 501, 'Z', 502, 'Z', 503, 'Z', 504, 'Z', 505, 'Z', 506, 'Z', 507, 'Z', 508, 'Z', 509, 'Z', 510, 'Z', 511, 'Z', 512, 'Z', 513, 'Z', 514, 'Z', 515, 'Z', 516, 'Z', 517, 'Z', 518, 'Z', 519, 'Z', 520, 'Z', 521, 'Z', 522, 'Z', 523, 'Z', 524, 'Z', 525, 'Z', 526, 'Z', 527, 'Z', 528, 'Z', 529, 'Z', 530, 'Z', 531, 'Z', 532, 'Z', 533, 'Z', 534, 'Z', 535, 'Z', 536, 'Z', 537, 'Z', 538, 'Z', 539, 'Z', 540, 'Z', 541, 'Z', 542, 'Z', 543, 'Z', 544, 'Z', 545, 'Z', 546, 'Z', 547, 'Z', 548, 'Z', 549, 'Z', 550, 'Z', 551, 'Z', 552, 'Z', 553, 'Z', 554, 'Z', 555, 'Z', 556, 'Z', 557, 'Z', 558, 'Z', 559, 'Z', 560, 'Z', 561, 'Z', 562, 'Z', 563, 'Z', 564, 'Z', 565, 'Z', 566, 'Z', 567, 'Z', 568, 'Z', 569, 'Z', 570, 'Z', 571, 'Z', 572, 'Z', 573, 'Z', 574, 'Z', 575, 'Z', 576, 'Z', 577, 'Z', 578, 'Z', 579, 'Z', 580, 'Z', 581, 'Z', 582, 'Z', 583, 'Z', 584, 'Z', 585, 'Z', 586, 'Z', 587, 'Z', 588, 'Z', 589, 'Z', 590, 'Z', 591, 'Z', 592, 'Z', 593, 'Z', 594, 'Z', 595, 'Z', 596, 'Z', 597, 'Z', 598, 'Z', 599, 'Z', 600, 'Z', 601, 'Z', 602, 'Z', 603, 'Z', 604, 'Z', 605, 'Z', 606, 'Z', 607, 'Z', 608, 'Z', 609, 'Z', 610, 'Z', 611, 'Z', 612, 'Z', 613, 'Z', 614, 'Z', 615, 'Z', 616, 'Z', 617, 'Z', 618, 'Z', 619, 'Z', 620, 'Z', 621, 'Z', 622, 'Z', 623, 'Z', 624, 'Z', 625, 'Z', 626, 'Z', 627, 'Z', 628, 'Z', 629, 'Z', 630, 'Z', 631, 'Z', 632, 'Z', 633, 'Z', 634, 'Z', 635, 'Z', 636, 'Z', 637, 'Z', 638, 'Z', 639, 'Z', 640, 'Z', 641, 'Z', 642, 'Z', 643, 'Z', 644, 'Z', 645, 'Z', 646, 'Z', 647, 'Z', 648, 'Z', 649, 'Z', 650, 'Z', 651, 'Z', 652, 'Z', 653, 'Z', 654, 'Z', 655, 'Z', 656, 'Z', 657, 'Z', 658, 'Z', 659, 'Z', 660, 'Z', 661, 'Z', 662, 'Z', 663, 'Z', 664, 'Z', 665, 'Z', 666, 'Z', 667, 'Z', 668, 'Z', 669, 'Z', 670, 'Z', 671, 'Z', 672, 'Z', 673, 'Z', 674, 'Z', 675, 'Z', 676, 'Z', 677, 'Z', 678, 'Z', 679, 'Z', 680, 'Z', 681, 'Z', 682, 'Z', 683, 'Z', 684, 'Z', 685, 'Z', 686, 'Z', 687, 'Z', 688, 'Z', 689, 'Z', 690, 'Z', 691, 'Z', 692, 'Z', 693, 'Z', 694, 'Z', 695, 'Z', 696, 'Z', 697, 'Z', 698, 'Z', 699, 'Z', 700, 'Z', 701, 'Z', 702, 'Z', 703, 'Z', 704, 'Z', 705, 'Z', 706, 'Z', 707, 'Z', 708, 'Z', 709, 'Z', 710, 'Z', 711, 'Z', 712, 'Z', 713, 'Z', 714, 'Z', 715, 'Z', 716, 'Z', 717, 'Z', 718, 'Z', 719, 'Z', 720, 'Z', 721, 'Z', 722, 'Z', 723, 'Z', 724, 'Z', 725, 'Z', 726, 'Z', 727, 'Z', 728, 'Z', 729, 'Z', 730, 'Z', 731, 'Z', 732, 'Z', 733, 'Z', 734, 'Z', 735, 'Z', 736, 'Z', 737, 'Z', 738, 'Z', 739, 'Z', 740, 'Z', 741, 'Z', 742, 'Z', 743, 'Z', 744, 'Z', 745, 'Z', 746, 'Z', 747, 'Z', 748, 'Z', 749, 'Z', 750, 'Z', 751, 'Z', 752, 'Z', 753, 'Z', 754, 'Z', 755, 'Z', 756, 'Z', 757, 'Z', 758, 'Z', 759, 'Z', 760, 'Z', 761, 'Z', 762, 'Z', 763, 'Z', 764, 'Z', 765, 'Z', 766, 'Z', 767, 'Z', 768, 'Z', 769, 'Z', 770, 'Z', 771, 'Z', 772, 'Z', 773, 'Z', 774, 'Z', 775, 'Z', 776, 'Z', 777, 'Z', 778, 'Z', 779, 'Z', 780, 'Z', 781, 'Z', 782, 'Z', 783, 'Z', 784, 'Z', 785, 'Z', 786, 'Z', 787, 'Z', 788, 'Z', 789, 'Z', 790, 'Z', 791, 'Z', 792, 'Z', 793, 'Z', 794, 'Z', 795, 'Z', 796, 'Z', 797, 'Z', 798, 'Z', 799, 'Z', 800, 'Z', 801, 'Z', 802, 'Z', 803, 'Z', 804, 'Z', 805, 'Z', 806, 'Z', 807, 'Z', 808, 'Z', 809, 'Z', 810, 'Z', 811, 'Z', 812, 'Z', 813, 'Z', 814, 'Z', 815, 'Z', 816, 'Z', 817, 'Z', 818, 'Z', 819, 'Z', 820, 'Z', 821, 'Z', 822, 'Z', 823, 'Z', 824, 'Z', 825, 'Z', 826, 'Z', 827, 'Z', 828, 'Z', 829, 'Z', 830, 'Z', 831, 'Z', 832, 'Z', 833, 'Z', 834, 'Z', 835, 'Z', 836, 'Z', 837, 'Z', 838, 'Z', 839, 'Z', 840, 'Z', 841, 'Z', 842, 'Z', 843, 'Z', 844, 'Z', 845, 'Z', 846, 'Z', 847, 'Z', 848, 'Z', 849, 'Z', 850, 'Z', 851, 'Z', 852, 'Z', 853, 'Z', 854, 'Z', 855, 'Z', 856, 'Z', 857, 'Z', 858, 'Z', 859, 'Z', 860, 'Z', 861, 'Z', 862, 'Z', 863, 'Z', 864, 'Z', 865, 'Z', 866, 'Z', 867, 'Z', 868, 'Z', 869, 'Z', 870, 'Z', 871, 'Z', 872, 'Z', 873, 'Z', 874, 'Z', 875, 'Z', 876, 'Z', 877, 'Z', 878, 'Z', 879, 'Z', 880, 'Z', 881, 'Z', 882, 'Z', 883, 'Z', 884, 'Z', 885, 'Z', 886, 'Z', 887, 'Z', 888, 'Z', 889, 'Z', 890, 'Z', 891, 'Z', 892, 'Z', 893, 'Z', 894, 'Z', 895, 'Z', 896, 'Z', 897, 'Z', 898, 'Z', 899, 'Z', 900, 'Z', 901, 'Z', 902, 'Z', 903, 'Z', 904, 'Z', 905, 'Z', 906, 'Z', 907, 'Z', 908, 'Z', 909, 'Z', 910, 'Z', 911, 'Z', 912, 'Z', 913, 'Z', 914, 'Z', 915, 'Z', 916, 'Z', 917, 'Z', 918, 'Z', 919, 'Z', 920, 'Z', 921, 'Z', 922, 'Z', 923, 'Z', 924, 'Z', 925, 'Z', 926, 'Z', 927, 'Z', 928, 'Z', 929, 'Z', 930, 'Z', 931, 'Z', 932, 'Z', 933, 'Z', 934, 'Z', 935, 'Z', 936, 'Z', 937, 'Z', 938, 'Z', 939, 'Z', 940, 'Z', 941, 'Z', 942, 'Z', 943, 'Z', 944, 'Z', 945, 'Z', 946, 'Z', 947, 'Z', 948, 'Z', 949, 'Z', 950, 'Z', 951, 'Z', 952, 'Z', 953, 'Z', 954, 'Z', 955, 'Z', 956, 'Z', 957, 'Z', 958, 'Z', 959, 'Z', 960, 'Z', 961, 'Z', 962, 'Z', 963, 'Z', 964, 'Z', 965, 'Z', 966, 'Z', 967, 'Z', 968, 'Z', 969, 'Z', 970, 'Z', 971, 'Z', 972, 'Z', 973, 'Z', 974, 'Z', 975, 'Z', 976, 'Z', 977, 'Z', 978, 'Z', 979, 'Z', 980, 'Z', 981, 'Z', 982, 'Z', 983, 'Z', 984, 'Z', 985, 'Z', 986, 'Z', 987, 'Z', 988, 'Z', 989, 'Z', 990, 'Z', 991, 'Z', 992, 'Z', 993, 'Z', 994, 'Z', 995, 'Z', 996, 'Z', 997, 'Z', 998, 'Z', 999, 'Z', 1000, 'Z', 1001, 'Z', 1002, 'Z', 1003, 'Z', 1004, 'Z', 1005, 'Z', 1006, 'Z', 1007, 'Z', 1008, 'Z', 1009, 'Z', 1010, 'Z', 1011, 'Z', 1012, 'Z', 1013, 'Z', 1014, 'Z', 1015, 'Z', 1016, 'Z', 1017, 'Z', 1018, 'Z', 1019, 'Z', 1020, 'Z', 1021, 'Z', 1022, 'Z', 1023, 'Z', 1024, 'Z', 1025, 'Z', 1026, 'Z', 1027, 'Z', 1028, 'Z', 1029, 'Z', 1030, 'Z', 1031, 'Z', 1032, 'Z', 1033, 'Z', 1034, 'Z', 1035, 'Z', 1036, 'Z', 1037, 'Z', 1038, 'Z', 1039, 'Z', 1040, 'Z', 1041, 'Z', 1042, 'Z', 1043, 'Z', 1044, 'Z', 1045, 'Z', 1046, 'Z', 1047, 'Z', 1048, 'Z', 1049, 'Z', 1050, 'Z', 1051, 'Z', 1052, 'Z', 1053, 'Z', 1054, 'Z', 1055, 'Z', 1056, 'Z', 1057, 'Z', 1058, 'Z', 1059, 'Z', 1060, 'Z', 1061, 'Z', 1062, 'Z', 1063, 'Z', 1064, 'Z', 1065, 'Z', 1066, 'Z', 1067, 'Z', 1068, 'Z', 1069, 'Z', 1070, 'Z', 1071, 'Z', 1072, 'Z', 1073, 'Z', 1074, 'Z', 1075, 'Z', 1076, 'Z', 1077, 'Z', 1078, 'Z', 1079, 'Z', 1080, 'Z', 1081, 'Z', 1082, 'Z', 1083, 'Z', 1084, 'Z', 1085, 'Z', 1086, 'Z', 1087, 'Z', 1088, 'Z', 1089, 'Z', 1090, 'Z', 1091, 'Z', 1092, 'Z', 1093, 'Z', 1094, 'Z', 1095, 'Z', 1096, 'Z', 1097, 'Z', 1098, 'Z', 1099, 'Z', 1100, 'Z', 1101, 'Z', 1102, 'Z', 1103, 'Z', 1104, 'Z', 1105, 'Z', 1106, 'Z', 1107, 'Z', 1108, 'Z', 1109, 'Z', 1110, 'Z', 1111, 'Z', 1112, 'Z', 1113, 'Z', 1114, 'Z', 1115, 'Z', 1116, 'Z', 1117, 'Z', 1118, 'Z', 1119, 'Z', 1120, 'Z', 1121, 'Z', 1122, 'Z', 1123, 'Z', 1124, 'Z', 1125, 'Z', 1126, 'Z', 1127, 'Z', 1128, 'Z', 1129, 'Z', 1130, 'Z', 1131, 'Z', 1132, 'Z', 1133, 'Z', 1134, 'Z', 1135, 'Z', 1136, 'Z', 1137, 'Z', 1138, 'Z', 1139, 'Z', 1140, 'Z', 1141, 'Z', 1142, 'Z', 1143, 'Z', 1144, 'Z', 1145, 'Z', 1146, 'Z', 1147, 'Z', 1148, 'Z', 1149, 'Z', 1150, 'Z', 1151, 'Z', 1152, 'Z', 1153, 'Z', 1154, 'Z', 1155, 'Z', 1156, 'Z', 1157, 'Z', 1158, 'Z', 1159, 'Z', 1160, 'Z', 1161, 'Z', 1162, 'Z', 1163, 'Z', 1164, 'Z', 1165, 'Z', 1166, 'Z', 1167, 'Z', 1168, 'Z', 1169, 'Z', 1170, 'Z', 1171, 'Z', 1172, 'Z', 1173, 'Z', 1174, 'Z', 1175, 'Z', 1176, 'Z', 1177, 'Z', 1178, 'Z', 1179, 'Z', 1180, 'Z', 1181, 'Z', 1182, 'Z', 1183, 'Z', 1184, 'Z', 1185, 'Z', 1186, 'Z', 1187, 'Z', 1188, 'Z', 1189, 'Z', 1190, 'Z', 1191, 'Z', 1192, 'Z', 1193, 'Z', 1194, 'Z', 1195, 'Z', 1196, 'Z', 1197, 'Z', 1198, 'Z', 1199, 'Z', 1200, 'Z', 1201, 'Z', 1202, 'Z', 1203, 'Z', 1204, 'Z', 1205, 'Z', 1206, 'Z', 1207, 'Z', 1208, 'Z', 1209, 'Z', 1210, 'Z', 1211, 'Z', 1212, 'Z', 1213, 'Z', 1214, 'Z', 1215, 'Z', 1216, 'Z', 1217, 'Z', 1218, 'Z', 1219, 'Z', 1220, 'Z', 1221, 'Z', 1222, 'Z', 1223, 'Z', 1224, 'Z', 1225, 'Z', 1226, 'Z', 1227, 'Z', 1228, 'Z', 1229, 'Z', 1230, 'Z', 1231, 'Z', 1232, 'Z', 1233, 'Z', 1234, 'Z', 1235, 'Z', 1236, 'Z', 1237, 'Z', 1238, 'Z', 1239, 'Z', 1240, 'Z', 1241, 'Z', 1242, 'Z', 1243, 'Z', 1244, 'Z', 1245, 'Z', 1246, 'Z', 1247, 'Z', 1248, 'Z', 1249, 'Z', 1250, 'Z', 1251, 'Z', 1252, 'Z', 1253, 'Z', 1254, 'Z', 1255, 'Z', 1256, 'Z', 1257, 'Z', 1258, 'Z', 1259, 'Z', 1260, 'Z', 1261, 'Z', 1262, 'Z', 1263, 'Z', 1264, 'Z', 1265, 'Z', 1266, 'Z', 1267, 'Z', 1268, 'Z', 1269, 'Z', 1270, 'Z', 1271, 'Z', 1272, 'Z', 1273, 'Z', 1274, 'Z', 1275, 'Z', 1276, 'Z', 1277, 'Z', 1278, 'Z', 1279, 'Z', 1280, 'Z', 1281, 'Z', 1282, 'Z', 1283, 'Z', 1284, 'Z', 1285, 'Z', 1286, 'Z', 1287, 'Z', 1288, 'Z', 1289, 'Z', 1290, 'Z', 1291, 'Z', 1292, 'Z', 1293, 'Z', 1294, 'Z', 1295, 'Z', 1296, 'Z', 1297, 'Z', 1298, 'Z', 1299, 'Z', 1300, 'Z', 1301, 'Z', 1302, 'Z', 1303, 'Z', 1304, 'Z', 1305, 'Z', 1306, 'Z', 1307, 'Z', 1308, 'Z', 1309, 'Z', 1310, 'Z', 1311, 'Z', 1312, 'Z', 1313, 'Z', 1314, 'Z', 1315, 'Z', 1316, 'Z', 1317, 'Z', 1318, 'Z', 1319, 'Z', 1320, 'Z', 1321, 'Z', 1322, 'Z', 1323, 'Z', 1324, 'Z', 1325, 'Z', 1326, 'Z', 1327, 'Z', 1328, 'Z', 1329, 'Z', 1330, 'Z', 1331, 'Z', 1332, 'Z', 1333, 'Z', 1334, 'Z', 1335, 'Z', 1336, 'Z', 1337, 'Z', 1338, 'Z', 1339, 'Z', 1340, 'Z', 1341, 'Z', 1342, 'Z', 1343, 'Z', 1344, 'Z', 1345, 'Z', 1346, 'Z', 1347, 'Z', 1348, 'Z', 1349, 'Z', 1350, 'Z', 1351, 'Z', 1352, 'Z', 1353, 'Z', 1354, 'Z', 1355, 'Z', 1356, 'Z', 1357, 'Z', 1358, 'Z', 1359, 'Z', 1360, 'Z', 1361, 'Z', 1362, 'Z', 1363, 'Z', 1364, 'Z', 1365, 'Z', 1366, 'Z', 1367, 'Z', 1368, 'Z', 1369, 'Z', 1370, 'Z', 1371, 'Z', 1372, 'Z', 1373, 'Z', 1374, 'Z', 1375, 'Z', 1376, 'Z', 1377, 'Z', 1378, 'Z', 1379, 'Z', 1380, 'Z', 1381, 'Z', 1382, 'Z', 1383, 'Z', 1384, 'Z', 1385, 'Z', 1386, 'Z', 1387, 'Z', 1388, 'Z', 1389, 'Z', 1390, 'Z', 1391, 'Z', 1392, 'Z', 1393, 'Z', 1394, 'Z', 1395, 'Z', 1396, 'Z', 1397, 'Z', 1398, 'Z', 1399, 'Z', 1400, 'Z', 1401, 'Z', 1402, 'Z', 1403, 'Z', 1404, 'Z', 1405, 'Z', 1406, 'Z', 1407, 'Z', 1408, 'Z', 1409, 'Z', 1410, 'Z', 1411, 'Z', 1412, 'Z', 1413, 'Z', 1414, 'Z', 1415, 'Z', 1416, 'Z', 1417, 'Z', 1418, 'Z', 1419, 'Z', 1420, 'Z', 1421, 'Z', 1422, 'Z', 1423, 'Z', 1424, 'Z', 1425, 'Z', 1426, 'Z', 1427, 'Z', 1428, 'Z', 1429, 'Z', 1430, 'Z', 1431, 'Z', 1432, 'Z', 1433, 'Z', 1434, 'Z', 1435, 'Z', 1436, 'Z', 1437, 'Z', 14

RESULT 14

S16366
collagen alpha 2(IV) chain precursor - pig roundworm
C:Species: Ascaris suum (pig roundworm)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S16366
R:Pattitt, J.; Kingston, I.B.
J. Biol. Chem. 266, 16149-16156, 1991
A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the partial cDNA sequence of the alpha-1(IV) chain.
A:Reference number: S16366; MUID:91340768; PMID:1714907
A:Accession: S16366
A:Molecule type: mRNA
A:Reids: 1-1763 <JBI>
A:Cross-references: UNIPROT:P27393; UNIPARC:UPI0000126D40; GB:M67507; NID:G159648; PIDN:C16366
C:Genetics:
A:Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1; 1011/1
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfide
E:1-26/Domain: signal sequence #status predicted <SIG>
F:27-1763/Product: collagen alpha 2(IV) chain #status predicted <SIG>
F:27-42/Domain: non-collagenous NH1 #status predicted <NH1>
F:43-1529/Domain: collagenous #status predicted <COL>
F:197-199/Region: cell attachment (R-G-D) motif
F:1530-1763/Domain: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F:1530-1638/Domain: repeat NC1 #status predicted <NC11>
F:1639-1763/Domain: repeat NC1 #status predicted <NC12>
F:31,34,39,41,536,539/Disulfide bonds: interchain #status predicted
F:126/Binding site: carboxylate (Asn) (covalent) #status predicted
F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 12.6%; Score 265; DB 2; Length 1763;
Best Local Similarity 34.3%; Pred. No. 1.6e-09;
Matches 85; Conservative 15; Mismatches 82; Indels 66; Gaps 10;

Qy	13	PAAAPRRSGOG-----CGCGAPAPARGSGNSCLLFLGFGLSLALHLTLCCYLELR	65
Db	256	PEGAPGMKGEKDPGEAGPRFPPTGVAGQP-----GLPGMKGERGL	298
Qy	66	SELRERGAESRLGGSTPGTSGTLSSLGGLDDPSITSHLGQSPKQOPLPGEAALHS	125
Db	299	SGPAGPRGKEGRPLGPPGPKGD-----RGLD-----GLPGVPLPGQKEAGF--	343
Qy	126	DSQDGHQWALLNFFFPDEKPYSEESRVRNRKRSKSNEGADGPVKNNKKKAGPPGN	185
Db	344	PGRDGAKGAR---GPGPPGGGSGFS-----DGPPGPGLPREGQPGPPGAD	387
Qy	186	GPPGPPGPPGPGQPGPIGPIGPIGTVMGPPGP-----PGPPGPGQPPGLQSPSG	236
Db	388	GYPGPPGPPGPGQPLGPGGGLPLGLPG--LEGLPGPKGEKGDSCIPGAPGVQPPGLAGPPG	445
Qy	237	AADKAGTR 244	
Db	446	AKGEPGPR 453	

RESULT 15

CGHUIB
collagen alpha 4(IV) chain precursor - human
N:Alternate names: procollagen alpha 4(IV) chain
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C:Accession: A55360; S38854; S38777
R:Leinonen, A.; Mariyama, M.; Mochizuki, T.; Tryggvason, K.; Reeders, S.T.
J. Biol. Chem. 269, 26172-26177, 1994
A:Title: Complete primary structure of the human type IV collagen alpha4(IV) chain. Complete primary structure of the human type IV collagen alpha4(IV) chain. Complete primary structure of the human type IV collagen alpha4(IV) chain. Complete primary structure of the human type IV collagen alpha4(IV) chain.
A:Reference number: A55360; MUID:95014445; PMID:7523402
A:Accession: A55360
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Reids: 1-1690 <LEI>
A:Cross-references: UNIPROT:P53420; UNIPARC:UPI0000000769; GB:X81053; NID:G574805; PIDN:C16366
R:Yoshimoto, M.; Ohashi, T.; Yoshioka, H.; Matsuo, N.; Ninomiya, Y.

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:21:42 ; Search time 298 Seconds
(without alignments)
1213.695 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRR.....NMSKHTTFGAIRLGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2098	100.0	391	1	EDA_HUMAN	Q92838 homo sapien
2	1972	94.0	391	1	EDA_BOVIN	Q9beg5 bos taurus
3	1965	93.7	391	1	EDA_MOUSE	O54693 mus musculus
4	1965	93.7	391	2	Q3UV69_MOUSE	Q3uv69 mus musculus
5	1890.5	90.1	386	2	Q5BP77_CANFA	Q5bp77 canis famill
6	1441.5	68.7	356	2	O5EFZ5_CHICK	Q5efz5 gallus gall
7	955.5	45.5	367	2	Q5OD54_GASAC	Q5od54 gasteroste
8	887	42.3	178	2	Q5J500_HUMAN	Q5j500 homo sapien
9	809	38.6	207	2	Q4SJRI_TETNG	Q4sjri tetraodon n
10	286	13.6	742	2	Q8VIF6_MOUSE	Q8vif6 mus musculus
11	284	13.5	998	2	Q8CFM4_MOUSE	Q8cfm4 mus musculus
12	282	13.4	742	2	Q8C979_MOUSE	Q8c979 mus musculus
13	282	13.4	742	2	Q8K4Q8_MOUSE	Q8k4q8 mus musculus
14	280.5	13.4	1792	2	Q59EE7_HUMAN	Q59ee7 homo sapien
15	280.5	13.4	1838	1	COSA1_HUMAN	P20908 homo sapien
16	280.5	13.4	1840	2	Q59IP3_PIG	Q59ip3 sus scrofa
17	280	13.3	284	2	Q25581_9BILA	Q25581 teladorsagi
18	278	13.3	742	2	Q4V985_RAT	Q4v985 rattus norv
19	277.5	13.2	1838	1	COSA1_MOUSE	O88207 mus musculus
20	277.5	13.2	1840	1	COSA1_RAT	Q9j103 rattus norv
21	276	13.2	622	2	Q9BY85_HUMAN	Q9by85 homo sapien
22	276	13.2	1315	2	Q8QHL9_XENLA	Q8qhl9 xenopus lae
23	275.5	13.1	1470	1	COHAI_MOUSE	Q07563 mus musculus
24	275	13.1	1049	1	CO3A1_BOVIN	P04258 bos taurus
25	275	13.1	1431	1	COHAI_MESAU	Q9jfm4 mesocricetu
26	273.5	13.0	1840	1	COSA1_CRILO	Q60467 cricetus
27	273	13.0	742	2	Q5KU26_HUMAN	O5ku26 homo sapien
28	273	13.0	742	2	Q8WZA4_HUMAN	O8wza4 homo sapien
29	273	13.0	1497	1	COHAI_HUMAN	Q9umd9 homo sapien
30	272	13.0	284	2	Q25582_9BILA	Q25582 teladorsagi
31	272	13.0	1415	2	Q5UV36_HUMAN	Q5uv36 homo sapien

32	272	13.0	1802	2	Q17163_BRUMA	Q17163 brugia mala
33	271.5	12.9	1344	2	O93419_CHICK	O93419 gallus gall
34	271	12.9	367	2	O5CIA3_SCHJA	O5cia3 schistosoma
35	270	12.9	1222	2	Q8K173_MOUSE	Q8k173 mus musculus
36	270	12.9	1464	1	CO3A1_MOUSE	.P08121 mus musculus
37	270	12.9	1464	2	Q3TVI5_MOUSE	Q3tvi5 m osteoclas
38	270	12.9	1464	2	Q3UH72_MOUSE	Q3uh72 mus musculus
39	270	12.9	1464	2	Q7TT32_MOUSE	Q7tt32 mus musculus
40	270	12.9	1464	2	Q8BLW4_MOUSE	Q8blw4 mus musculus
41	270	12.9	1467	2	Q5DTG2_MOUSE	Q5dtg2 mus musculus
42	268.5	12.8	742	2	Q6P9F2_HUMAN	Q6p9f2 homo sapien
43	268.5	12.8	742	2	Q9BYH7_HUMAN	Q9byh7 homo sapien
44	268.5	12.8	1258	2	Q8AW11_BRARE	Q8aw11 brachydanio
45	268.5	12.8	1835	2	Q9IAU4_CHICK	Q9iau4 gallus gall

ALIGNMENTS

RESULT 1

ID EDA_HUMAN STANDARD; PRT; 391 AA.
AC Q92838; O75910; Q5JUM7; Q9Y6L0; Q9Y6L1; Q9Y6L2; Q9Y6L3;
AC Q9Y6L4;
DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
DT 15-JUL-1999, sequence version 2.
DT 07-FEB-2006, entry version 53.
DE Ectodysplasin-A (Ectodermal dysplasia protein) (EDA protein)
DE [Contains: Ectodysplasin-A, membrane form; Ectodysplasin-A, secreted form].
GN Name:EDA; Synonyms=ED1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] (ISOFORM I), AND VARIANTS EDA HIS-61 AND LEU-69.
RC TISSUE=Sweat gland;
RX MEDLINE=96331280; PubMed=8696334;
RA Kere J., Srivastava A.K., Montonen O., Zonana J., Thomas N.S.T., Ferguson B.M., Munoz F., Morgan D., Clarke A., Baybayan P., Chen E.Y., Ezer S., Saarialho-Kere U., la Chapelle A., Schlessinger D.;
RA "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein.";
RL Nat. Genet. 13:409-416(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM A1), AND VARIANTS EDA. TISSUE=Liver;
RC MEDLINE=98349961; PubMed=9683615;
RA Monreal A.W., Zonana J., Ferguson B.M.;
RA "Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations.";
RL Am. J. Hum. Genet. 63:380-389(1998).
RN [3]
RP NUCLEOTIDE SEQUENCE (ISOFORMS A1; A2; B; C; D; E AND F), AND VARIANTS EDA. MEDLINE=98409495; PubMed=9736768; DOI=10.1093/hmg/7.11.1661;
RA Bayes M., Hartung A.J., Ezer S., Pispas J., Thesleff I., Srivastava A.K., Kere J.;
RA "The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative splicing and encodes ectodysplasin-A with deletion mutations in collagenous repeats.";
RL Hum. Mol. Genet. 7:1661-1669(1998).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15772651; DOI=10.1038/nature03440;
RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McEay K., Muzny D., Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A., Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Wen G., Jones M.C., Hurler M.E., Andrews T.D., Scott C.E., Searle S.,

RA Rameer J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzker M.L.,
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,
RA Worley K.C., Ainscough R., Ambrose K.D., Ansari-Lari M.A., Aradhyia S.,
RA Ashwell R.I., Babarge A.K., Baguley C.P., Ballabio A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgman A.M., Brown A.J., Brown M.J., Bonnin D.,
RA Bruford E.A., Buhay C., Burch P., Burford D., Burgess J., Burrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinnault C.,
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blanchet K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Deshazo D., Dhani P., Ding Y., Dinh H., Dodsworth S.,
RA Eades T., Ellwood M., Emery-Cohen A., Errington H., Evans K.L.,
RA Faulkner L., Francis F., Frankland J., Fraser A.E., Galgoczy P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heitmann K., Hennig S., Hernandez J.,
RA Hinzmann B., Ho S., Hoffs M., Howden P.J., Huckle E.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kioschis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,
RA Lloyd C., Lloyd D.M., Loulsegod H., Loveland J.E., Lovell J.D.,
RA Lozado R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,
RA McDowell J., McLaren S., McMurray A., Meidl P., Meitinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordieck G., Nyakatura G., O'dell C.N.,
RA Okwuonu G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pellan S.E.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,
RA Ridler K.A., Schlesinger D., Schuler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Shownkeen R., Skuce C.D.,
RA Smith M.L., Sotheman E.C., Steingruber H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbreck D., Taber P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Urso M., Verdusco D., Villaseña D., Waldron L.,
RA Wall M., Wang Q., Warren J., Warr G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.J., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williams A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,
RA Lehrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.,
RA "The DNA sequence of the human X chromosome.";
RL Nature 434:325-337(2005).
RL [5]
RN TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
RP Kobiela K., Kobiela A., Trzciak W.H.;
RT "Expression of a novel transcript isoform of the EDA gene in human
RT umbilical cord.";
RL Eur. J. Hum. Genet. Suppl. 7:104-104(1999).
RL [6]
RN RECEPTOR INTERACTION (ISOFORMS A1 AND A2).
RX MEDLINE=20495245; PubMed=11039935; DOI=10.1126/science.290.5491.523;
RA Yan M., Wang L.-C., Hymowitz S.G., Schillbach S., Lee J., Goddard A.,
RA de Vos A.M., Gao W.-Q., Dixit V.M.;
RT "Two-amino acid molecular switch in an epithelial morphogen that
RT regulates binding to two distinct receptors.";
RL Science 290:523-527(2000).
RN [7]
RN PROCESSING, MUTAGENESIS OF ARG-153, AND CHARACTERIZATION OF VARIANT
RP HIS-156.
RX MEDLINE=21205766; PubMed=11309369; DOI=10.1093/bmg/10.9.953;
RA Elomaa O., Pulkkinen K., Hanneilus U., Mikkola M., Saarialho-Kere U.,
RA Kere J.;
RT "Ectodysplasin is released by proteolytic shedding and binds to the
RT EDAR protein.";

RL Hum. Mol. Genet. 10:953-962(2001).
RN [8]
RN CHARACTERIZATION OF VARIANTS CYS-155; CYS-156 AND HIS-156, MUTAGENESIS
RP OF ARG-153; LYS-158 AND ARG-159, AND CLEAVAGE SITE.
RX MEDLINE=21309955; PubMed=11416205; DOI=10.1073/pnas.131076098;
RA Chen Y., Molloy S.S., Thomas L., Gambee J., Baechinger H.P.,
RA Ferguson B.M., Zonana J., Thomas G., Morris N.P.;
RT "Mutations within a furin consensus sequence block proteolytic release
RT of ectodysplasin-A and cause X-linked hypohidrotic ectodermal
RT dysplasia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7218-7223(2001).
RN [9]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 230-391, AND SUBUNIT.
RX PubMed=14656435; DOI=10.1016/j.str.2003.11.009;
RA Hymowitz S.G., Compain D.M., Yan M., Wallweber H.J., Dixit V.M.,
RA Starovasnik M.A., de Vos A.M.;
RT "The crystal structures of EDA-A1 and EDA-A2: splice variants with
RT distinct receptor specificity.";
RL Structure 11:1513-1520(2003).
RN [10]
RN VARIANT EDA TYR-54.
RX MEDLINE=98292028; PubMed=9630076;
RA Hertz J.M., Noergaard Hansen K., Juncker I., Kjeldsen M.,
RA Gregersen N.;
RT "A novel missense mutation (402C-->T) in exon 1 in the EDA gene in a
RT family with X-linked hypohidrotic ectodermal dysplasia.";
RL Clin. Genet. 53:205-209(1998).
RN [11]
RN VARIANT EDA LYS-63.
RX MEDLINE=98168231; PubMed=9507389;
RA Ferguson B.M., Thomas N.S.T., Munoz F., Morgan D., Clarke A.,
RA Zonana J.;
RT "Scarcity of mutations detected in families with X linked hypohidrotic
RT ectodermal dysplasia: diagnostic implications.";
RL J. Med. Genet. 35:112-115(1998).
RN [12]
RN VARIANT EDA ARG-55.
RX MEDLINE=99393307; PubMed=10469321;
RX DOI=10.1046/j.1523-1747.1999.00656.x;
RA Martinez F., Millan J.M., Orellana C., Prieto F.;
RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia caused by a
RT novel mutation in EDAL gene: 406T > G (Leu55Arg).";
RL J. Invest. Dermatol. 113:285-286(1999).
RN [13]
RN VARIANTS EDA ARG-60; TYR-252; VAL-269; SER-302 AND MET-378.
RX MEDLINE=21272350; PubMed=11378824; DOI=10.1038/sj.ehbg/5200635;
RA Vincent M.C., Biancalana V., Giniesty D., Mandel J.-L., Calvas P.;
RT "Mutational spectrum of the ED1 gene in X-linked hypohidrotic
RT ectodermal dysplasia.";
RL Eur. J. Hum. Genet. 9:355-363(2001).
RN [14]
RN VARIANTS EDA CYS-156; HIS-156; CYS-255; ASP-255; GLY-274; TYR-332 AND
RP THR-349.
RX MEDLINE=21193173; PubMed=11295832; DOI=10.1002/humu.33;
RA Paekkoenen K., Cambiaghi S., Novelli G., Ouzts L.V., Penttinen M.,
RA Kere J., Srivastava A.K.;
RT "The mutation spectrum of the EDA gene in X-linked anhidrotic
RT ectodermal dysplasia.";
RL Hum. Mutat. 17:349-349(2001).
CC -!- FUNCTION: Seems to be involved in epithelial-mesenchymal signaling
CC during morphogenesis of ectodermal organs. Isoform A1 binds only
CC to the receptor EDAR, while isoform A2 binds exclusively to the
CC receptor XEDAR.
CC -!- SUBUNIT: Homotrimer. The homotrimers may then dimerize and form
CC higher order oligomers.
CC -!- INTERACTION:
CC QUNEO:EDAR; NbExp=1; IntAct=EBI-529425, EBI-529289;
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=8;
CC Comment=Additional isoforms seem to exist;
CC Name=A1; Synonyms=I1;
CC IsoId=Q92838-1; Sequence=Displayed;

Db
361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 3

ID	EDA MOUSE	STANDARD;	PRT;	391 AA.
AC	Q54693;	Q5705; Q9QWJ8; Q9QZ01;	Q9QZ02;	
CD	15-JUL-1999,	integrated into UniProtKB/Swiss-Prot.		
DT	15-JUL-1999,	sequence version 2.		
DT	07-FEB-2006,	entry version 44.		
DE	Ectodysplasin-A (EDA protein homolog) (Tabby protein) [Contains:			
DE	Ectodysplasin-A, membrane form; Ectodysplasin-A, secreted form].			
GN	Name=Eda; Synonyms=Edl, Ta;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muroidae; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RP	NUCLEOTIDE SEQUENCE (ISOFORMS TAA; TAB AND TAC).			
RC	STRAIN=129/Sv;			
RC	MEDLINE=98058770; PubMed=9371801; DOI=10.1073/pnas.94.24.13069;			
RA	Strivastava A.K., Pispis J., Hartung A.J., Du Y., Ezer S., Jenks T.,			
RA	Shimada T., Pekkanen M., Mikkola M.L., Ko M.S.H., Theislev I.,			
RA	Kere J., Schlesinger D.;			
RT	"The Tabby phenotype is caused by mutation in a mouse homologue of the			
RT	Eda gene that reveals novel mouse and human exons and encodes a			
RT	protein (ectodysplasin-A) with collagenous domains.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 94:13069-13074(1997).			
RP	[2]			
RP	NUCLEOTIDE SEQUENCE (ISOFORM TAD).			
RC	MEDLINE=97449184; PubMed=9285798; DOI=10.1093/hmg/6.9.1589;			
RC	Ferguson B.M., Brockdorff N., Formstone E., Ngyuen T.,			
RA	Krommiller J.E., Zonana J.;			
RA	"Cloning of Tabby, the murine homolog of the human EDA gene: evidence			
RT	for a membrane-associated protein with a short collagenous domain.";			
RL	Hum. Mol. Genet. 6:1589-1594(1997).			
RP	[3]			
RP	NUCLEOTIDE SEQUENCE (ISOFORMS TA-A2 AND TA-A3).			
RC	TISSUE=Embryo;			
RC	MEDLINE=20005791; PubMed=10534613; DOI=10.1016/S0925-4773(99)00180-X;			
RA	Mikkola M.L., Pispis J., Pekkanen M., Paulin L., Nieminen P., Kere J.,			
RA	Theislev I.;			
RT	"Ectodysplasin, a protein required for epithelial morphogenesis, is a			
RT	novel TNF homologue and promotes cell-matrix adhesion.";			
RL	Mech. Dev. 88:133-146(1999).			
CC	-1- FUNCTION: Involved in epithelial-mesenchymal signaling during			
CC	morphogenesis of ectodermal organs. Isoform TAA binds only to the			
CC	receptor EDAR, while isoform TA-A2 binds exclusively to the			
CC	receptor XEDAR.			
CC	-1- SUBUNIT: Homotrimer (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By			
CC	similarity).			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event-alternative splicing; Named isoforms=6;			
CC	Comment=Additional isoforms seem to exist;			
CC	Name=TAA; Synonyms=A1;			
CC	IsoId=O54693-1; Sequence=Displayed;			
CC	Name=TA-A2;			
CC	IsoId=O54693-2; Sequence=VSP_006471;			
CC	Name=TA-A3;			
CC	IsoId=O54693-3; Sequence=VSP_006469, VSP_006471;			
CC	Name=TAB;			
CC	IsoId=O54693-4; Sequence=VSP_006466, VSP_006467;			
CC	Name=TAC;			
CC	IsoId=O54693-5; Sequence=VSP_006465, VSP_006468;			
CC	Name=TAD;			
CC	IsoId=O54693-6; Sequence=VSP_006470;			
CC	-1- PTM: N-glycosylated (By similarity).			
CC	-1- PTM: Processing by furin produces a secreted form (By similarity)			
CC	-1- DISEASE: Defects in Eda are the cause of the tabby phenotype in			
CC	mice (the equivalent of anhidrotic ectodermal dysplasia in			

	humans). The disease is characterized by sparse hair (atrachiosis or hypotrichosis), abnormal or missing teeth and the inability to sweat due to the absence of sweat glands.
-	SIMILARITY: Belongs to the tumor necrosis factor family.
-	SIMILARITY: Contains 1 collagen-like domain.
<hr/>	
CC	Copysighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC	Distributed under the Creative Commons Attribution-NoDerivs License
<hr/>	
DR	ENBL; AF016627; AAB95202.1; -; mRNA.
DR	ENBL; AF016628; AAB95203.1; -; mRNA.
DR	ENBL; AF016629; AAB95204.1; -; mRNA.
DR	ENBL; AF016630; AAB95205.1; -; mRNA.
DR	ENBL; AF016631; AAB95206.1; -; mRNA.
DR	ENBL; AF004434; AAB88121.1; -; Genomic_DNA.
DR	ENBL; AF004435; AAB88122.1; -; mRNA.
DR	ENBL; Y13438; CAA73849.1; -; Genomic_DNA.
DR	ENBL; AJ243657; CAB52696.1; -; mRNA.
DR	ENBL; AJ243658; CAB52697.1; -; mRNA.
DR	SNR; O54693; 242-390.
DR	Ensembl; ENSMUSG0000059327; Mus musculus.
DR	MGI; MGI:1195272; Eda.
DR	GO; GO:0045177; C.apical part of cell; IDA.
DR	GO; GO:0005789; C.endoplasmic reticulum membrane; IDA.
DR	GO; GO:0005987; C.integral to plasma membrane; IDA.
DR	GO; GO:0007160; P.cell-matrix adhesion; IDA.
DR	GO; GO:0042346; P.positive regulation of NF-kappaB-nucleus im. . ; IDA.
DR	GO; GO:0007431; P.salivary gland development; IDA.
DR	InterPro; IPR008160; Collagen.
DR	InterPro; IPR006052; TNF_family.
DR	Pfam; PF01391; Collagen; 1.
DR	Pfam; PF00229; TNF; 1.
DR	PROSITE; PS00251; TNF_1; FALSE_NEG.
DR	PROSITE; PS50049; TNF_2; 1.
KW	Alternative splicing; Collagen; Developmental protein;
KW	Differentiation; Glycoprotein; Membrane; Signal-anchor; Transmembrane.
FT	CHAIN 1 391 Ectodysplasin-A, membrane form.
FT	CHAIN 160 391 Ectodysplasin-A, secreted form (By similarity).
FT	/FTId=PRO_0000034540.
FT	/FTId=PRO_0000034541.
FT	Cytoplasmic (Potential).
FT	Signal-anchor for type II membrane protein (Potential).
FT	Extracellular (Potential).
FT	Collagen-like.
FT	Cleavage (by furin) (By similarity).
FT	N-linked (GlcNac. .) (Potential).
FT	N-linked (GlcNac. .) (Potential).
FT	MALLNFPPDFKAYSEESRRVRRNRKSKSGADGFVRNK
FT	KKGKAGPDPNGPGPPGPDPGGPGIPGIGPTVTM
FT	GPSPGPPGPGPGPLQGPSGA -> VSHLGGAAALEAP
FT	SPARLGGLSLRAQGTLPURAKFPQRSWSHGVLRGCPCGQ
FT	VFLSGCLGSRPVPMSWAQAAPAAPGEAWAA (in isoform TAC).
FT	/FTId=VSP_006465.
FT	/FTId=VSP_006466.
FT	PVKNNKKKK -> KSTQVIFFP (in isoform TAB).
FT	/FTId=VSP_006466.
FT	Missing (in isoform TAB).
FT	/FTId=VSP_006467.
FT	Missing (in isoform TAC).
FT	/FTId=VSP_006468.
FT	Missing (in isoform TA-A3).
FT	/FTId=VSP_006469.
FT	Missing (in isoform TAD).
FT	/FTId=VSP_006470.
FT	Missing (in isoform TA-A2 and isoform TA-A3).
FT	/FTId=VSP_006471.
FT	D >> E (in Ref. 2).
FT	CONFLICT 126 126
FT	SEQUENCE 391 AA; 41603 MW; E5CEDA5BD6DEFF CRC64;
SQ	Query Match 93.7%; Score 1965; DB 1; Length 391;

Query Match


```
SQ SEQUENCE 386 AA; 41235 MW; 6095BDBC3EACF68B CRC64;
Query Match 90.1%; Score 1890.5; DB 2; Length 386;
Best Local Similarity 91.3%; Pred. No. 1.9e-108;
Matches 357; Conservative 8; Mismatches 21; Indels 5; Gaps 2;
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGSGNSCLLFLGFLSLALHLTLCC 60
DB 1 MGYPEVERREPLTAAPRERGSGCGCGGAPAPAGSGNSCLLFLGFLSLALHLTLCC 60
QY 61 YLELRSELRRERGAERLGGSGTPTGTSGLTSSLGGLDPPSPITSHLGQSPKQOPLPGE 120
DB 61 YLELRSELRRERGAERLGGSGTPTGTSGLTSSLGGLDPPSPITSHLGQSPKQOPLPGE 120
QY 121 AALHSDSQDHQWALLNFFPDPKPYSEESRRVRNKRKSKSNEGADGVPVNNKKKKKAG 180
DB 117 TALPHSRDCHQWALLNFFPDPKPYSEESRRVRNKRKSKSNEGADGVPVNNKKKKKAG 180
QY 181 PCPNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 176 PCPNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGPRENQPAVVHLQGGSAIQVKNLDSGGVLDWSRITWNPVKVFKLHPRSGEVLVDGT 300
DB 236 AGPRENQPAVVHLQGGSAIQVKNLDSGGVLDWSRITWNPVKVFKLHPRSGEVLVDGT 300
QY 301 YFTYSQVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCTYAGVCLLKARQKI 360
DB 296 YFTYSQVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCTYAGVCLLKARQKI 360
QY 361 AVKQVHADISINNSKHTTFFGAILRGEAPAS 391
DB 356 AVKQVHADISINNSKHTTFFGAILRGEAPAS 386

RESULT 6
QSEFZ5 CHICK
ID QSEFZ5_CHICK PRELIMINARY; PRT; 356 AA.
AC QSEFZ5;
DT 15-MAR-2005, integrated into UniProtKB/TREMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Ectodysplasin A1 (Fragment).
GN Name=Eda;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Houghton L., Morgan B.A.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC Science 307:1928-1933(2005).
DR EMBL; AY885699; AAW81692.1; -; mRNA.
DR SMR; O5EFZ5; 207-355.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:000164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR006052; TNF_family.
DR Pfam; PF01391; Collagen; 2.
DR DR PROSITE; PS00229; TNF; 1.
DR DR PROSITE; PS50049; TNF_2; 1.
DR NON_TER 1
SQ SEQUENCE 356 AA; 37527 MW; 8D53FD97F71FDDFB CRC64;
Query Match 68.7%; Score 1441.5; DB 2; Length 356;
Best Local Similarity 74.9%; Pred. No. 7.7e-81;
Matches 289; Conservative 19; Mismatches 41; Indels 37; Gaps 9;
QY 13 PAAAPRERSQ-----CGCGGAPAPAGSGNSCLLFLGFLSLALHLTLCCYLELRSEL 68
DB 1 PAAMGAERREGEAARGCCSCAG-----GSMFLFLGFFALSALAHVLTLCYLELRSEL 53
QY 69 RREERGAESHLGGSGTPTGTSGLTSSLGGLDPPSPITSHLGQSPK-QOPLPGEAALHSDS 127
DB 54 RRDGRGQP-----AAPRRDGTAAAAA---PGAP-----PAVRPQPAESGERR----- 94
QY 128 QDHQWALLNFFPDPKPYSEESRRVRNKRKSKSNEGADGVPVNNKKKKKAGPPGP 165
DB 95 -----QQALLNFFPPEKJH-VGEGRVRNKRKSKSGEGDPGSSVKNKKKKKAGPPGP 150
QY 186 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 245
DB 151 GQGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 210
QY 246 NPAPVHLQGGSAIQVKNLDSGGVLDWSRITWNPVKVFKLHPRSGEVLVDGTTFIYS 305
DB 211 TQPAVHLQGGSAIQVKNLDSGGVLDWSRITWNPVKVFKLHPRSGEVLVDGTTFIYS 270
QY 306 QVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCTYAGVCLLKARQKI 365
DB 271 QVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCTYAGVCLLKARQKI 330
QY 366 HADISINNSKHTTFFGAILRGEAPAS 391
DB 331 HADISINNSKHTTFFGAILRGEAPAS 356

RESULT 7
Q50D54 GASAC
ID Q50D54_GASAC PRELIMINARY; PRT; 367 AA.
AC Q50D54;
DT 07-JUN-2005, integrated into UniProtKB/TREMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Ectodysplasin.
GN Name=Eda;
OS Gasterosteus aculeatus (Three-spined stickleback).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
OC Gasterosteidae; Gasterosteus.
OX NCBI_TaxID=69293;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA PubMed=15790847; DOI=10.1126/science.1107239;
RA Colosimo P.F., Hosemann K.E., Balabhadra S., Villarreal G. Jr.,
RA Dickson M., Hromadka J., Schmutz J., Myers R.M., Schluter D.,
RA Kingsley D.M.;
RT Widespread parallel evolution in sticklebacks by repeated fixation of
RT Ectodysplasin alleles.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC Science 307:1928-1933(2005).
DR EMBL; AY897589; AAY27076.1; -; Genomic DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:000164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR006052; TNF_family.
DR Pfam; PF01391; Collagen; 1.
DR DR PROSITE; PS00229; TNF; 1.
DR DR PROSITE; PS00251; TNF_1; UNKNOWN_1.
DR PROSITE; PS00049; TNF_2; 1.
```



```
KW Collagen.
SQ SEQUENCE 367 AA; 39746 MW; C0BPEC40D53A9EF6 CRC64;

Query Match 45.5%; Score 955.5; DB 2; Length 367;
Best Local Similarity 51.1%; Pred. No. 6.5e-51;
Matches 192; Conservative 59; Mismatches 86; Indels 39; Gaps 9;

QY 25 CCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCCYLELRSELRERGAESRLGGSGTP 84
Db 19 CTC-NKKKRSRGS--VFLGLFLLSLHAVTLVCLLRSEVRKEIHKR----- 68
QY 85 GTSGLTSSLGGLDPSDPSITSHL-GQSPKQOPLPEGAEALHSDSQDGHQMALHFFFPDE 143
Db 69 ---DTMLTLAGIDPADPAALAPQPRDPDGVGRGSE-----GHEEKLH---RNG 113
QY 144 KPYSEESRVRNRKSKSNEGADGPV-----KXKKGKAGPDPGNGPPGPPGPG 195
Db 114 ELHATQDNRG1--TORAKRSPGQPETELTGRERRKKEKKKRPVP---GPPGPPGPG 168
QY 196 PQGPPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPG 255
Db 169 PQGPPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPG 228
QY 256 QGSAIQVNDLSGVLNDSRITMNPVKFLHPRSGEVLVDGTYFYISQVEVYVYINF 315
Db 229 QETTIQVREDLSGILRNKMWVSIHRVFKHRSRSGEVLVDGTYFYISQVEVYVYINF 288
QY 316 DPASYEVVVDKPFLOCTSIETGKNTYNTCTAGVCLLKARQKIAVMVHADISINMSK 375
Db 289 DIASYEVVVDNPFLLCTSIETGKNTCTAGVCLLKARQKISIRIVYEDTLISMTN 348
QY 376 HTTFCAIRLGEAPAS 391
Db 349 HTFLGSLVRLGEAPSA 364

RESULT 8
Q5J500_HUMAN PRELIMINARY; PRT; 178 AA.
ID Q5J500
AC Q5J500
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Ectodysplasin A.
GN Name=EDA; ORFNames=RP11-351X23.1-003;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP NUCLEOTIDE SEQUENCE.
RA Whitehead S.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RA Clark S.;
RP NUCLEOTIDE SEQUENCE.
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; AL450449; CAI39805.1; -; Genomic_DNA.
DR EMBL; AL158069; CAI39805.1; JOINED; Genomic_DNA.
DR EMBL; AL158069; CAI41611.1; -; Genomic_DNA.
DR EMBL; AL450449; CAI41611.1; JOINED; Genomic_DNA.
DR Ensembl; ENSG00000158813; Homo sapiens.
SQ SEQUENCE 178 AA; 19200 MW; A390EF5B3F6B865A CRC64;

Query Match 42.3%; Score 887; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 5e-47;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MGYPEVERRELLPAAAPRRERGSGCGCGCAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db 1 MGYPEVERRELLPAAAPRRERGSGCGCGCAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELRSELRERGAESRLGGSGTPTGTSGLSLGGLDPSDPSITSHLGGSPKQOPLPEGE 120
Db 61 YLELRSELRERGAESRLGGSGTPTGTSGLSLGGLDPSDPSITSHLGGSPKQOPLPEGE 120
QY 121 AALHSDSQDGHQMALHFFFPDEKPYSEESRVRNRKSKSNEGADG 168
Db 121 AALHSDSQDGHQMALHFFFPDEKPYSEESRVRNRKSKSNEGADG 168

RESULT 9
Q4SJRL_TETNG PRELIMINARY; PRT; 207 AA.
ID Q4SJRL_TETNG
AC Q4SJRL;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Chromosome 1 SCAP14573, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00017090001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lucifalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Cattolico L., Poullain J., De Berardinis V.,
RA Biemont C., Skalli Z., Cattolico L., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier S., Brottier P., Coutanceau J.-P., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
DR EMBL; CAAB01014573; CAF99121.1; -; Genomic_DNA.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0016020; Membrane; IEA.
DR GO; GO:0005164; F-tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P-immune response; IEA.
DR GO; GO:0006817; P-phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR006052; TNF family.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00229; TNF; 1.
DR PROSITE; PS00251; TNF 1; UNKNOWN_1.
DR PROSITE; PS50049; TNF_2; 1.
KW Collagen.
FT NON_TER 1
FT NON_TER 207
SQ SEQUENCE 207 AA; 22149 MW; 6CB874AF177DA5CA CRC64;
```



```

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC013626; AAH13626.1; -; mRNA.
DR Ensembl; ENSMUSG0000026043; Mus musculus.
DR MGI; MGI:88453; Col3a1.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0006817; P-phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 12.
DR ProDom; PD000007; Clg_helix; 2.
KW Collagen.
FT NON_TER
SQ SEQUENCE 998 AA; 95451 MW; 2AB4BA953B7084A6 CRC64;
Query Match 13.5%; Score 284; DB 2; Length 998;
Best Local Similarity 30.3%; Pred. No. 4e-09;
Matches 87; Conservative 15; Mismatches 91; Indels 94; Gaps 11;
QY 13 PAAAPRERSQG-----CGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCCYLELRSEL 68
DB 548 PRGGGGERGEGHPPGPPGAPGPPGAPGQNGEPGA----- 578
QY 69 RBERGAESRLGSGTGTCTGTSGLSLGLDPDPSITSHLGQSPKQOPLRPGERALHSDSQ 128
DB 579 KGERGAPGKGEHPPGPPGAPGPTGSSGAPGPPG-----QGVKGRSGPFGPGTAG----- 628
QY 129 DGHQMALNFFPFDPKPYSEESRRVRNRKRSKSGADGPPVKNKKKGKAGPPGPPGPP 188
DB 629 -----FP-----GGRGLGPPGN-----NGNPPGPPGSGAP 654
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 284
DB 655 GKDPGPPGAGNSGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 714
QY 246 NQPAV-----VHLOQGSATQVND-----LSGGVLDWMSRITWNPVKF 284
DB 715 GPPGPPGPPRVALDLRVSRRVKNQEPVAINGNV-----VLLDPKVF 755
RESULT 12
Q8C979_MOUSE PRELIMINARY; PRT; 742 AA.
AC Q8C979;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DE 07-FEB-2006, entry version 23.
DE 7 days neonate cerebellum cDNA, RIKEN full-length enriched library,
DE clone: A730023E20 product: collectin sub-family member 12, full insert
DE sequence.
GN Name=Collect12;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells K., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
Ambei-Imbimbato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
Banaal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crowe M.L., Dalla E., Dairymple B.F., de Bono B., Della Gatta G.,
di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Humiński L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liu S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,
Mortagui-Tabar S., Mulder N., Nakano N., Nakauchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
Tamaoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawahashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Plesey C., Shibata K., Shiraki T., Suzuki S.,
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome."
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome."
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oasato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran M.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayashizaki N.,
RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

```


DB 969 GHPGPPGPEQGLPLGKAGCTK-GDPGAPGLPGK 1003

RESULT 15

COSAL HUMAN

ID COSAL HUMAN STANDARD; PRT; 1838 AA.

AC 203008; Q15094; Q5SUX4;

DT 01-FEB-1991, integrated into UniProtKB/Swiss-Prot.

DT 13-SEP-2005, sequence version 3.

DT 07-MAR-2006, entry version 71.

DE Collagen alpha-1(V) chain precursor.

GN Name=COL5A1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN [1]_TaxID=9606;

RN [1]_TaxID=9606;

RP NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 556-565.

RX MEDLINE=91302336; PubMed=2071595;

RA Takahara K., Seto Y., Okasawa K., Okamoto N., Noda A., Yaoi Y.,

RA Kato I.;

RT "Complete primary structure of human collagen alpha 1 (V) chain.";

RL J. Biol. Chem. 266:13124-13129(1991).

RN [2]

RP NUCLEOTIDE SEQUENCE [MRNA].

RX MEDLINE=92105142; PubMed=1722213;

RA Greenspan D.S., Cheng W., Hoffman G.G.;

RT "The pro-alpha-1(V) collagen chain: complete primary structure,

RT distribution of expression, and comparison with the pro-alpha-1(XI)

RT collagen chain.";

RL J. Biol. Chem. 266:24727-24733(1991).

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX PubMed=15164053; DOI=10.1038/nature02465;

RA Humphray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E.,

RA Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,

RA Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,

RA Babbage A.K., Babbage S., Baggeley C.L., Bailey J., Banerjee R.,

RA Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,

RA Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burtill W.,

RA Burton J., Carder C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,

RA Clark S.Y., Clee C.M., Clegg S., Collier R.E., Corby N., Crosier M.,

RA Cummings A.T., Davies J., Dhali P., Dunn M., Dutta I., Dyer L.W.,

RA Earthrowl M.E., Faulkner L., Fleming C.J., Garner P., Garnett J.,

RA Frankland J.A., French L., Fricker D.G., Frankish A.,

RA Ghorji J., Gilbert J.G.R., Gilson C., Graham D.V., Gribble S.,

RA Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,

RA Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,

RA Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,

RA Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,

RA Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,

RA Lawlor S., Leongamornlert D.A., Leverhulme M., Lloyd C., Lloyd D.M.,

RA Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,

RA McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,

RA Norddick G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,

RA Peltan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,

RA Senda H.K., Showkeen R., Sims S.K., Skuce C.D., Smith M.,

RA Steward C.A., Swarbreck D., Sycamore N., Tester J., Thorpe A.,

RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,

RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,

RA Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,

RA Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,

RA Rogers J., Dunham I.;

RT "DNA sequence and analysis of human chromosome 9.";

RL Nature 429:369-374(2004).

RN [4]

RP PROTEIN SEQUENCE OF 621-822.

RX TISSUE=Chorioamniotic membrane;

RC MEDLINE=89227189; PubMed=2496661;

RA Seyer J.M., Kang A.H.;

RT "Covalent structure of collagen: amino acid sequence of three cyanogen bromide-derived peptides from human alpha 1(V) collagen chain.";

RL Arch. Biochem. Biophys. 271:120-129(1989).

[5]

RP PROTEIN SEQUENCE OF 823-950, AND HEPARIN-BINDING.

RX MEDLINE=90366601; PubMed=2203476; DOI=10.1016/0304-4165(90)90108-9;

RA Yaoi Y., Hashimoto K., Koitabashi H., Takahara K., Ito M., Kato I.;

RT "Primary structure of the heparin-binding site of type V collagen.";

RL Biochim. Biophys. Acta 1035:139-145(1990).

RN [6]

RP PROTEIN SEQUENCE OF 556-571.

RC TISSUE=Placenta;

RX MEDLINE=92239022; PubMed=15711108;

RA Mann K.;

RT "Isolation of the alpha 3-chain of human type V collagen and

RT characterization by partial sequencing.";

RL Biol. Chem. Hoppe-Seyler 373:69-75(1992).

RN [7]

RP PROTEIN SEQUENCE OF 565-576; 756-772; 1012-1029; 1219-1232 AND

RP 1465-1477.

RC TISSUE=Chorioamniotic membrane;

RX MEDLINE=94237164; PubMed=8181482;

RA Moradi-Ameli M., Rousseau J.C., Kleman J.P., Champlaud M.F.,

RA Bouillon M.M., Bernillon J., Wallach J.M., van der Rest M.;

RT "Diversity in the processing events at the N-terminus of type-V

RT collagen.";

RL Eur. J. Biochem. 221:987-995(1994).

RN [8]

RP INTERACTION WITH CSPG4.

RX PubMed=9099729; DOI=10.1074/jbc.272.16.10769;

RA Tillet E., Ruggiero F., Nishiyama A., Stallcup W.B.;

RT "The membrane-spanning proteoglycan NG2 binds to collagens V and VI

RT through the central nonglobular domain of its core protein.";

RL J. Biol. Chem. 272:10769-10776(1997).

RN [9]

RP DISEASE, AND VARIANT EDS-I SER-1639.

RX MEDLINE=97195540; PubMed=9042913;

RA de Paeppe A., Nuytinck L., Hausser I., Anton-Lamprecht I.,

RA Naeyaert J.-M.;

RT "Mutations in the COL5A1 gene are causal in the Ehlers-Danlos

RT syndromes I and II.";

RL Am. J. Hum. Genet. 60:547-554(1997).

RN [10]

RP VARIANTS EDS-1 SER-530 AND ASP-1489.

RX MEDLINE=20068401; PubMed=10602121;

RA DOI=10.1002/(SICI)1096-8628(20000103)90:1<72::AID-AJMG13>3.0.CO;2-C;

RA Giunta C., Steinmann B.;

RT "Compound heterozygosity for a disease-causing G1489D and disease-

RT modifying G530S substitution in COL5A1 of a patient with the classical

RT type of Ehlers-Danlos syndrome: an explanation of intrafamilial

RT variability?";

RL Am. J. Med. Genet. 90:72-79(2000).

CC -!- FUNCTION: Type V collagen is a member of group I collagen

CC (fibrillar forming collagen). It is a minor connective tissue

CC component of nearly ubiquitous distribution. Type V collagen binds

CC to DNA, heparan sulfate, thrombospondin, heparin, and insulin.

CC -!- SUBUNIT: Trimers of two alpha 1(V) and one alpha 2(V) chains in

CC most tissues and trimers of one alpha 1(V), one alpha 2(V), and

CC one alpha 3(V) chains in placenta. Interacts with CSPG4.

CC -!- PTM: Prolines at the third position of the tripeptide repeating

CC unit (G-X-Y) are hydroxylated in some or all of the chains.

CC -!- PTM: Sulfated on 40% of tyrosines

CC -!- DISEASE: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome

CC type I (EDS-I) [MIM:130000]; also known as Ehlers-Danlos

CC syndrome mitis. Inheritance is autosomal dominant.

CC -!- SIMILARITY: Belongs to the fibrillar collagen family.

CC -!- SIMILARITY: Contains 1 laminin G-like domain.

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; D90279; BAA14323.1; -; mRNA.
DR EMBL; M76729; AAS9993.1; -; mRNA.
DR EMBL; AL591890; CAI15483.1; -; Genomic DNA.
DR EMBL; AL603650; CAI15483.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI15483.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI17261.1; -; Genomic DNA.
DR EMBL; AL591890; CAI17261.1; JOINED; Genomic DNA.
DR EMBL; AL603650; CAI17261.1; JOINED; Genomic DNA.
DR EMBL; AL603650; CAI39859.1; -; Genomic DNA.
DR EMBL; AL591890; CAI39859.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI39859.1; JOINED; Genomic DNA.
DR PIR; S18802; CGHUIV.
DR PDB; 1A89; Model; A/B/C=904-924.
DR PDB; 1A9A; Model; A/C=904-924.
DR Ensembl; ENSG00000130635; Homo sapiens.
DR HGNC; HGNC:2209; COL5A1.
DR MIM; 120215; Gene.
DR MIM; 130000; Phenotype.
DR MIM; 130010; Phenotype.
DR GO; GO:0005588; C:collagen type V; TAS.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR Pfam; PF02210; Laminin_G_2; 1.
DR ProDom; PD000007; Clg_helix; 4.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS50025; LAM_G_DOMAIN; FALSE NEG.

KW 3D-structure; Collagen; Direct protein sequencing; Disease mutation;
KW Ehlers-Danlos syndrome; Extracellular matrix; Heparin-binding;
KW Hydroxylation; Repeat; Signal; Structural protein; Sulfation.
FT SIGNAL 1 37 Potential.
FT CHAIN 38 1605 Collagen alpha-1(V) chain.
FT PROPEP 1606 1838 /FTid=PRO_0000005756.
FT C-terminal propeptide. /FTid=PRO_0000005757.
FT TSP N-terminal.
FT DOMAIN 39 230 Laminin G-like.
FT REGION 231 443 Nonhelical region.
FT REGION 444 558 Interrupted collagenous region.
FT REGION 559 1570 Triple-helical region.
FT REGION 1571 1605 Nonhelical region.
FT MOD_RES 234 234 Sulfotyrosine (Potential).
FT MOD_RES 236 236 Sulfotyrosine (Potential).
FT MOD_RES 240 240 Sulfotyrosine (Potential).
FT MOD_RES 262 262 Sulfotyrosine (Potential).
FT MOD_RES 263 263 Sulfotyrosine (Potential).
FT MOD_RES 338 338 Sulfotyrosine (Potential).
FT MOD_RES 340 340 Sulfotyrosine (Potential).

Search completed: June 2, 2006, 17:30:03
Job time : 301 secs

Query Match 13.4%; Score 280.5; DB 1; Length 1838;
Best Local Similarity 30.1%; Pred. No. 1.2e-08;
Matches 83; Conservative 9; Mismatches 69; Indels 115; Gaps 10;
QY 13 PAAAPRERSQCG-GCGGAPARAGENSCLF.GFGFLSLALHLLTLCYLELRSELRE 71
Db 857 PLGPPGEGKGLGVPLGYPGRQPKGS-----IGFPG:P----- 891
QY 72 RGNESRLGSGTGTSGTSLSSGLGLDPDPSITSHLQSPSKOPLPEGEALHSDSQGH 131
Db 892 -GANGKGGRGTP-----GKPGRQGRPTG----- 916
QY 132 QMALLNFFFPDEKPYSEESRRVRNRKRSNEGADGPVKNKKKGKAGPP---GPNGPP 188

Db 917 -----PRGERGPRGITKPGPKGNSGSDGP-----AGPPGERGPNGPQ 954
QY 189 GPPGPPGQGGPGIGIPGIPG-----TTVMGPPGPPGP----- 222
Db 955 GPTGPPGPKGPPGPPGKDLPGHPGQRGETGFGKTPGPPGVVGPQGTGTGPMGR 1014
QY 223 --PGPQGGPPGLQGPSGAADKAGTRENQPAVVHLQQQ 256
Db 1015 GHPPGPPGPGEGQLPLGKAGKGTK-GDRCPAGLPGK 1049

[illegible]

QY 1 MGYPEVERRELLPAAAPRRSGCGCGAPARAGSGNCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPEVERREPLPAAAPRRSGCGCGRGPAPARAGSGNCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRERGAESRLGSGTPTGTLSSLGGLDPSPTITSHLGQSPKQOPLPGE 120
DB 61 YLELSELRERGTESRLGSGAPGTGTLSSPGSLDPVGPITRHLGQSPKQOPLPGE 120
QY 121 AALHSDSQDHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAG 180
DB 121 DFLPPESQRHQWALLNFFPDEKAYSEESRRVRNRKSKSNEGADGPNKKNKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQVKNLDSGGVLDWSRITMNPVKFLHPRSGELVLVDGT 300
DB 241 TGTRENQPAVHLQGGSAIQVKNLDSGGVLDWSRITMNPVKFLHPRSGEL- 293
QY 301 YFIYSQVEVYINFTDFASVEVVDKPFLOCTRSIETGKNTNTCTAGVCLLKARQKI 360
DB 294 -----EYIYNFTDFASVEVVDKPFLOCTRSIETGKNTNTCTAGVCLLKARQKI 346
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
DB 347 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 377
RESULT 5
US-09-949-016-11119
; Sequence 1119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11119
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11119
Query Match 66.9%; Score 1403; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.2e-110;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 132 QWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPGPPGPP 191
DB 1 QWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPGPPGPP 60
QY 192 GPPGQGGPPGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 251
DB 61 GPPGQGGPPGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPI 120
QY 252 HLOGGSAIQVKNLDSGGVLDWSRITMNPVKFLHPRSGELVLVDGTFTYISQVEVY 311
DB 121 HLOGGSAIQVKNLDSGGVLDWSRITMNPVKFLHPRSGELVLVDGTFTYISQVEVY 180
QY 312 INFTDFASVEVVDKPFLOCTRSIETGKNTNTCTAGVCLLKARQKI 371
DB 181 INFTDFASVEVVDKPFLOCTRSIETGKNTNTCTAGVCLLKARQKI 240

QY 372 NMSKHTTFFGAIRLGEAPAS 391
DB 241 NMSKHTTFFGAIRLGEAPAS 260
RESULT 6
US-07-609-716-66
; Sequence 66, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1889
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-716-66
Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;
QY 71 ERGAESRLGSGTPTGTLSSLGGLDPSPTITSHLGQSPKQOPLPGEAALHSDSQ 128
DB 71 DRGDAGPKGADGSPGAPGVPVSGAPGPPGPPGPPGPPGPPGPPGPPGPPG 130
QY 129 DGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPG 188
DB 131 D-----RGDAGPKGADGSPGAPGVPVSGAPGPPGPPGPPGPP 165
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 212
DB 166 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 225
QY 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 244
DB 226 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 258
RESULT 7
US-08-642-255-33
; Sequence 33, Application US/08642255
; Patent No. 5773249

```

; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-642-255-33

```

```

Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGASRLGGSTGTTSLSLGLDPS--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADSGPGAGPGVGSFGAPGPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQWALLNFFPFDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAGPPGNGPP 188
Db 131 D-----RGDAGPKGADSGPGAGPGVGSFGAPGPGPGPGPGPGPG 165
QY 189 GPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 212
Db 166 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 225
QY 213 -VMGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 244
Db 226 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 258

```

```

RESULT 8
US-08-475-411A-66
; Sequence 66, Application US/08475411A
; Patent No. 6140072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Flehr, Hobach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400

```

```

; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,411A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Tregartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-411A-66

```

```

Query Match 14.0%; Score 294; DB 2; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGASRLGGSTGTTSLSLGLDPS--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADSGPGAGPGVGSFGAPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQWALLNFFPFDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAGPPGNGPP 188
Db 131 D-----RGDAGPKGADSGPGAGPGVGSFGAPGPGPGPGPGPG 165
QY 189 GPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 212
Db 166 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 225
QY 213 -VMGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 244
Db 226 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 258

```

```

RESULT 9
US-08-478-029A-66
; Sequence 66, Application US/08478029A
; Patent No. 6184348
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:

```

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,029A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/609,716
FILING DATE: 06-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/269,429
FILING DATE: 09-NOV-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/114,618
FILING DATE: 29-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 06/927,258
FILING DATE: 04-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard P.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 66:

SEQUENCE CHARACTERISTICS:

LENGTH: 357 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-478-029A-66

Query Match 14.0%; Score 294; DB 2; Length 357;

Best Local Similarity 32.4%; Pred. No. 1.2e-16;

Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGASRLGGSTGTSLSLGLDPPSP--ITSHLQPSKQKQPLEPGEAALHSDSQ 128

DB 71 DRGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 130

QY 129 DGHQALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPKVKKKKKAGPGPGNGPP 188

DB 131 D-----RGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPGPG 165

QY 189 GPPG 212

DB 166 GPPGAGPG 225

QY 213 -VWGPPG 244

DB 226 GPPGAGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 258

RESULT 10

US-08-642-255-32

Sequence 32, Application US/08642255

Patent No. 5773249

GENERAL INFORMATION:

APPLICANT: CAPPELLO, Joseph

APPLICANT: FERRARI, Franco A.

TITLE OF INVENTION: High Molecular Weight Collagen-Like

TITLE OF INVENTION: Protein Polymers

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

NUMBER OF SEQUENCES: 135
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,255
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: ROWLAND, Bertram I.
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A55556-3/BIR
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LENGTH: 330 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-642-255-32

Query Match 13.2%; Score 277; DB 1; Length 330;

Best Local Similarity 39.4%; Pred. No. 2.9e-15;

Matches 65; Conservative 5; Mismatches 61; Indels 34; Gaps 4;

QY 79 GSGTPTGTSLSLGLDPPSPITSHLQPSKQKQPLEPGEAALHSDSQDGHQMALNLF 138

DB 58 GPPGPGPGAGPVGSPGAGPGPGP---PGPPPGPGAGPGPGP----- 94

QY 139 FFPDEKPYSEESRRVRNRKSKSNEGADGPKVKKKKKAGPGPGNGPPGPGPGPG 198

DB 95 --PPGPGPGPGAGPV-----GSPGAGPGPGPGPGPGPGPGPGPGPGPG 145

QY 199 PGIGPGICPGCTTVMGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 243

DB 146 PVGSPGAGPGPGPP--GPPPGPGAGPGPGPGPGPGPGPGPGPGPG 188

RESULT 11

US-07-609-716-65

Sequence 65, Application US/07609716

Patent No. 5514581

GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.

APPLICANT: Cappello, Joseph

TITLE OF INVENTION: Functional Recombinantly Prepared

TITLE OF INVENTION: Synthetic Protein Polymer

NUMBER OF SEQUENCES: 118

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: US

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30


```

; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-478-029A-65

Query Match 13.2%; Score 277; DB 2; Length 408;
Best Local Similarity 39.4%; Pred. No. 3.9e-15;
Matches 65; Conservative 5; Mismatches 6; Indels 34; Gaps 4;

Qy 79 GSGTGTGTGTLSSGLDPSPTITSHLQSPKQPLEPGEALHSDSQDHQWALLNF 138
Db 58 GPPGPPGAGVSCAGPPGP-----PGPPGPGAGPPG----- 94

Qy 139 FFDKPYSEESRRVRNRKSKSNEGADGVNKKKKGAGPPGPPGPPGPPGQ 198
Db 95 --PPGPPGPPGAGPV-----GSPGAPGPPGPPGPPGPPGPPGPPG 145

Qy 199 PGGPGIGPGTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 243
Db 146 PVGSPGAPGPPGPP--GPPGPPGAPGPPGPPGPPGPPGPPGPPGPPG 188

RESULT 14
US-09-710-239-25
; Sequence 25, Application US/09710239
; Patent No. 6992172
; GENERAL INFORMATION:
; APPLICANT: Chang, Robert C.
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; TITLE OF INVENTION: RECOMBINANT GELATINS
; FILE REFERENCE: FG0219
; CURRENT APPLICATION NUMBER: US/09/710,239
; CURRENT FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 416
; TYPE: PRT
; ORGANISM: human
; US-09-710-239-25

Query Match 12.7%; Score 265.5; DB 2; Length 416;
Best Local Similarity 29.7%; Pred. No. 3.7e-14;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRERSQG-----CGCGGAPARAGE-----GNSCLFLGFFGLSLHLTLCCY 61
Db 160 PAGAPDGKESGSPGAPGTGARGAPGDRGEPGPPGA-----GFAG----- 201

Qy 62 LEIIRSELRRERGAESRLGSGTGTSLSSGLDPSPTITSHLQSPKQPLEPGEA 121
Db 202 -----PPGADGQPGAKGEPGADGAKGADGAPGPPGAPG-----AGPPGPGNVGAPGA- 247

Qy 122 ALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGVNKKKKGAGP 181
Db 248 -----KGARSGAPPGATGPPGAAGRVGP 271

Qy 182 PGNGPPGPPGPPG-----QGPGIGIPGIPGTVMGPPGPPGPPGPPGPPGPPG 235
Db 182 PGNGPPGPPGPPG-----QGPGIGIPGIPGTVMGPPGPPGPPGPPGPPGPPG 235

; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-478-029A-65

Query Match 13.2%; Score 277; DB 2; Length 408;
Best Local Similarity 39.4%; Pred. No. 3.9e-15;
Matches 65; Conservative 5; Mismatches 6; Indels 34; Gaps 4;

Qy 79 GSGTGTGTGTLSSGLDPSPTITSHLQSPKQPLEPGEALHSDSQDHQWALLNF 138
Db 58 GPPGPPGAGVSCAGPPGP-----PGPPGPPGAGPPG----- 94

Qy 139 FFDKPYSEESRRVRNRKSKSNEGADGVNKKKKGAGPPGPPGPPGPPGQ 198
Db 95 --PPGPPGPPGAGPV-----GSPGAPGPPGPPGPPGPPGPPGPPG 145

Qy 199 PGGPGIGPGTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 243
Db 146 PVGSPGAPGPPGPP--GPPGPPGAPGPPGPPGPPGPPGPPGPPGPPG 188

RESULT 14
US-09-710-239-25
; Sequence 25, Application US/09710239
; Patent No. 6992172
; GENERAL INFORMATION:
; APPLICANT: Chang, Robert C.
; APPLICANT: Kivirikko, Kari I.
; APPLICANT: Neff, Thomas B.
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; TITLE OF INVENTION: RECOMBINANT GELATINS
; FILE REFERENCE: FG0219
; CURRENT APPLICATION NUMBER: US/09/710,239
; CURRENT FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 416
; TYPE: PRT
; ORGANISM: human
; US-09-710-239-25

Query Match 12.7%; Score 265.5; DB 2; Length 416;
Best Local Similarity 29.7%; Pred. No. 3.7e-14;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

Qy 13 PAAAPRERSQG-----CGCGGAPARAGE-----GNSCLFLGFFGLSLHLTLCCY 61
Db 160 PAGAPDGKESGSPGAPGTGARGAPGDRGEPGPPGA-----GFAG----- 201

Qy 62 LEIIRSELRRERGAESRLGSGTGTSLSSGLDPSPTITSHLQSPKQPLEPGEA 121
Db 202 -----PPGADGQPGAKGEPGADGAKGADGAPGPPGAPG-----AGPPGPGNVGAPGA- 247

Qy 122 ALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGVNKKKKGAGP 181
Db 248 -----KGARSGAPPGATGPPGAAGRVGP 271

Qy 182 PGNGPPGPPGPPG-----QGPGIGIPGIPGTVMGPPGPPGPPGPPGPPGPPG 235
Db 182 PGNGPPGPPGPPG-----QGPGIGIPGIPGTVMGPPGPPGPPGPPGPPGPPG 235

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:42:30 ; Search time 187 Seconds
(without alignments) 968.540 Million cell u

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRER.....NMSKHTTFFGAIRLGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs. 463214858 residues

Total number of hits satisfying chosen parameters: 20977797

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match of

Fast Processing: Minimum Match 0%
Maximum Match 100%

Maximum March 1996
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 3: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US09_PUBCOMB.pcp.*
- 4: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 5: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 6: /EMC_Celerra_SIDS3/prodata/2/pubpaa/US11_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2098	100.0	391	3	US-09-813-329-8	Sequence 8, Appli
2	2098	100.0	391	3	US-09-729-658B-2	Sequence 2, Appli
3	2098	100.0	391	4	US-10-218-547-42	Sequence 42, Appli
4	2098	100.0	391	4	US-10-012-605C-6	Sequence 6, Appli
5	2098	100.0	391	4	US-10-310-793-46	Sequence 46, Appli
6	2098	100.0	391	4	US-10-202-063-42	Sequence 42, Appli
7	2098	100.0	391	4	US-10-719-205-6	Sequence 6, Appli
8	2098	100.0	391	6	US-11-028-780-42	Sequence 42, Appli
9	2098	100.0	391	6	US-11-142-736-8	Sequence 8, Appli
10	1965	93.7	391	3	US-09-813-329-9	Sequence 9, Appli
11	1965	93.7	391	6	US-11-142-736-9	Sequence 9, Appli
12	1873	89.3	377	3	US-09-729-658B-4	Sequence 4, Appli
13	779	37.1	412	5	US-10-503-999-6	Sequence 6, Appli
14	759	36.2	410	5	US-10-503-999-8	Sequence 8, Appli
15	542	25.8	246	4	US-10-185-425-9	Sequence 9, Appli
16	542	25.8	246	5	US-10-871-775-9	Sequence 9, Appli
17	286	13.6	742	5	US-10-820-155-10	Sequence 10, Appli
18	282	13.4	742	4	US-10-203-860-4	Sequence 4, Appli
19	282	13.4	742	5	US-10-820-155-2	Sequence 2, Appli
20	282	13.4	742	5	US-10-820-155-6	Sequence 6, Appli
21	280.5	13.4	1838	5	US-10-852-335A-162	Sequence 162, App
22	280.5	13.4	1838	5	US-10-287-436A-464	Sequence 464, App
23	280.5	13.4	1838	5	US-10-287-436A-1161	Sequence 1161, Ap
24	280.5	13.4	1874	5	US-10-821-234-1182	Sequence 1182, Ap
25	276	13.2	622	5	US-10-820-155-9	Sequence 9, Appli
26	275.5	13.1	180	5	US-10-797-606-39	Sequence 39, Appli
27	275	13.1	1466	4	US-10-402-089-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-813-329-8

; Sequence 8, Application US/09813329

; Patent No. US20020012968A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Suibb Company

ATTORNEY: LITTON MYERS BURNS COMPANY
: TITLE OF INVENTION: No. US20020012968A1e1

TITLE OF INVENTION:	NO. 03200200123000
TITLE OF INVENTION:	Variants Thereof

; TITLE OF INVENTION: VALVE
: FILE REFERENCE: D0016.DD

FILE REFERENCE: D0018.IIP
CURRENT APPLICATION NUMBER: US/09/813.329

; CURRENT APPLICATION NUMBER: US/0
 : CURRENT FILING DATE: 2001-03-20

; CURRENT FILING DATE: 2001-03-20
: PRIOR APPLICATION NUMBER: 60/190 816

;; PRIOR APPLICATION NUMBER: 60/
: PRIOR FILING DATE: 2000-03-21

; PRIOR FILING DATE: 2000-0
: NUMBER OF SEQ ID NOS: 65

```

; NUMBER OF SEQ ID NOS: 65
: SOFTWARE: Biotin version 3.0

```

; SOFTWARE:

; SEQ ID NO 8

```

; LENGTH: 391

```

```

; TYPE: PRT

```

; ORGANISM: D

Query Match	100.0%	Score 2098;	DB 3;	Length 391;
Best Local Similarity	100.0%;	Pred. No. 1.6e-137;		
Matches 391;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```
|||||
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 2
US-09-729-658B-2
; Sequence 2, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729, 658B
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-658B-2

Query Match 100.0%; Score 2098; DB 3; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSSGLGDPSPITSHLGQSPKQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSSGLGDPSPITSHLGQSPKQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNLSSGVLNDWSRITWNPKVKLHPRSGELEVLDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNLSSGVLNDWSRITWNPKVKLHPRSGELEVLDGT 300
QY 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 3
US-10-218-547-42
; Sequence 42, Application US/10218547
; Publication No. US20030100074A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Methods And Compositions For Treating Metabolic Bone Diseases Rel
; FILE REFERENCE: PF561
; CURRENT APPLICATION NUMBER: US/10/218, 547
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/312,542
```

```
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/330,761
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 391
; TYPE: PRT
; ORGANISM: human
US-10-218-547-42

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSSGLGDPSPITSHLGQSPKQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGGSGTPTGTSGLSSGLGDPSPITSHLGQSPKQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNLSSGVLNDWSRITWNPKVKLHPRSGELEVLDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNLSSGVLNDWSRITWNPKVKLHPRSGELEVLDGT 300
QY 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 4
US-10-012-605C-6
; Sequence 6, Application US/10012605C
; Publication No. US20030104602A1
; GENERAL INFORMATION:
; APPLICANT: Haldeman, Betty A.
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE REFERENCE: 00-111
; CURRENT APPLICATION NUMBER: US/10/012,605C
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/254,019
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-605C-6

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGEGNSCLLPLGFFGLSLALHLLTLCC 60
```


[illegible]

RESULT 5
 US-10-310-793-46
 ; Sequence 46, Application US/10310793
 ; Publication No. US20030198640A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Yu, Guo-Liang
 ; APPLICANT: Ni, Jian
 ; APPLICANT: Rosen, Craig A
 ; APPLICANT: Zhang, Jun
 ; APPLICANT: Wei, Ping
 ; TITLE OF INVENTION: Methods And Compositions For Treating Inflammatory Bowel Diseases
 ; TITLE OF INVENTION: Relating To Human Tumor Necrosis Factor-Gamma Beta
 ; FILE REFERENCE: PF573
 ; CURRENT APPLICATION NUMBER: US/10/310,793
 ; CURRENT FILING DATE: 2002-12-06
 ; PRIOR APPLICATION NUMBER: 60/336,695
 ; PRIOR FILING DATE: 2001-12-07
 ; PRIOR APPLICATION NUMBER: 10/226,294
 ; PRIOR FILING DATE: 2002-08-23
 ; PRIOR APPLICATION NUMBER: 60/314,381
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 09/899,059
 ; PRIOR FILING DATE: 2001-07-06
 ; PRIOR APPLICATION NUMBER: 60/278,449
 ; PRIOR FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/216,879
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 09/559,290
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 60/180,908
 ; PRIOR FILING DATE: 2000-02-08
 ; PRIOR APPLICATION NUMBER: 60/134,067
 ; PRIOR FILING DATE: 1999-05-13
 ; PRIOR APPLICATION NUMBER: 60/132,227
 ; PRIOR FILING DATE: 1999-05-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 46
 ; LENGTH: 391
 ; TYPE: PRT
 ; ORGANISM: human
 US-10-310-793-46

[illegible]

Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPKVFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 7
US-10-719-205-6
; Sequence 6, Application US/10719205
; Publication No. US20040086971A1
; GENERAL INFORMATION:
; APPLICANT: Haldeman, Betty A.
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE OF INVENTION: ZACRP3X2
; FILE REFERENCE: 00-111
; CURRENT APPLICATION NUMBER: US/10/719,205
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/012,605C
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/254,019
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-205-6

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQOQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQOQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
QY 181 PRGPNPQPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PRGPNPQPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPKVFKLHPRSGEVLVDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPKVFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 8
US-11-028-780-42
; Sequence 42, Application US/11028780
; Publication No. US20050163747A1

; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; FILE REFERENCE: PFS59C1
; CURRENT APPLICATION NUMBER: US/11/028,780
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 10/202,062
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,838
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 42
; LENGTH: 391
; TYPE: PRT
; ORGANISM: human
US-11-028-780-42

Query Match 100.0%; Score 2098; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQOQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGSGTPTGTSGLSLGGLDPSDPTSHLGQSPKQOQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
QY 181 PRGPNPQPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PRGPNPQPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPKVFKLHPRSGEVLVDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPKVFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDPASVEVVVDKPFLOCTRSIETGKTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 9
US-11-142-736-8
; Sequence 8, Application US/11142736
; Publication No. US20050227283A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNF") ar
; FILE REFERENCE: D0016 DIV1
; CURRENT APPLICATION NUMBER: US/11/142,736
; CURRENT FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-142-736-8

Query Match 100.0%; Score 2098; DB 6; Length 391;


```

; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Mus musculus
;
US-09-729-658B-4

Query Match      89.3%; Score 1873; DB 3; Length 377;
Best Local Similarity 90.8%; Pred. No. 6.4e-122;
Matches 355; Conservative 2; Mismatches 20; Indels 14; Gaps 1;

QY 1 MGYPEVERRELPAAAPRRGSGCGCGGAPAPRAGSGNSCLLFLGPFGLSLALHLTLCC 60
DB 1 MGYPEVERRELPAAAPRRGSGCGCGGAPAPRAGSGNSCLLFLGPFGLSLALHLTLCC 60
QY 61 YLELRSELRRGAESRLGSGTPTGTSGLTSSLGGLDPSDPSITSHLGQSPKQOPLRGE 120
DB 61 YLELRSELRRGTERSLGSGGAPGSGTSSPSGLDPVGPITRHLGQSPFOOQPLRGE 120
QY 121 AALHSDSQDHOMALLNFFPDEKPYSEBSRRVRNRKRSKNEGADGPVNKKKGGKAG 180
DB 121 DPLPPSQDQHOMALLNFFPDEKAYSEBSRRVRNRKRSKSGEGADGPVNKKKGGKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIOVKNDLSGGVLDNWSRITWNPVKVFLHPRSGEVLVDGT 300
DB 241 TGTRENQPAVHLQGGSAIOVKNDLSGGVLDNWSRITWNPVKVFLHPRSGEL----- 293
QY 301 YFTYSQVEVYINFTDFASVEVVDKPFLOQCTRSIETGKTNVNTCYTAGVCLLKARQKI 360
DB 294 -----EYVYINFTDFASVEVVDKPFLOQCTRSIETGKTNVNTCYTAGVCLLKARQKI 346
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
DB 347 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 377

RESULT 13
US-10-503-999-6
; Sequence 6, Application US/10503999
; Publication No. US20050152872A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; PRIOR FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-8

Query Match      36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.7e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
DB 242 GKRSPQPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPP 295
QY 275 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQVEVYINFTDFASVEVVDKPFLOQCTR 334
DB 296 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQ--VYINFTDFASVEVVDKPFLOQCTR 353
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 391
DB 354 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:
;

```

```

; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-6

Query Match      37.1%; Score 779; DB 5; Length 412;
Best Local Similarity 87.6%; Pred. No. 7.1e-46;
Matches 155; Conservative 1; Mismatches 15; Indels 6; Gaps 2;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
DB 242 GKRSPQPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPP 295
QY 275 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQVEVYINFTDFASVEVVDKPFLOQCTR 334
DB 296 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQVEVYINFTDFASVEVVDKPFLOQCTR 355
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 391
DB 356 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 412

RESULT 14
US-10-503-999-8
; Sequence 8, Application US/10503999
; Publication No. US20050152872A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; PRIOR FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: PCT/EP02/09354
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR APPLICATION NUMBER: DE 102 05 583.1
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 410
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-8

Query Match      36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.7e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 274
DB 242 GKRSPQPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPPKPP 295
QY 275 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQVEVYINFTDFASVEVVDKPFLOQCTR 334
DB 296 SRITWNPVKVFLHPRSGEVLVDGTFTYFSQ--VYINFTDFASVEVVDKPFLOQCTR 353
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 391
DB 354 SIETGKTNVNTCYTAGVCLLKARQKIAVKMWHADISINMSKHTTFFGAIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:
;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: June 2, 2006, 17:42:48 ; Search time 15 Seconds
 (without alignments)
 301.465 Million cell updates/sec

Title: US-09-729-658C-2
 Perfect score: 2098
 Sequence: 1 MGYPVERRELLPAAAPRER.....NMSKHTTFGAIRLGEAPAS 391

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA New.*
 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB pep.*
 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB pep.*
 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB pep.*
 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB pep.*
 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB pep.*
 7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB pep.*
 8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	265.5	12.7	1464	6 US-10-985-570-2	Sequence 2, Appli
2	263	12.5	1366	6 US-10-985-570-3	Sequence 3, Appli
3	256	12.2	1466	6 US-10-985-570-1	Sequence 1, Appli
4	252	12.0	744	6 US-10-505-928-493	Sequence 493, App
5	250	11.9	717	6 US-10-505-928-438	Sequence 438, App
6	244.5	11.7	314	7 US-11-293-697-3733	Sequence 3733, App
7	220.5	10.5	520	6 US-10-196-743-332	Sequence 332, App
8	186.5	8.9	399	6 US-10-953-349-22288	Sequence 22288, A
9	184	8.8	37	6 US-10-499-266-22	Sequence 22, Appl
10	150	7.1	542	7 US-11-293-697-4713	Sequence 4713, App
11	146	7.0	1085	6 US-10-505-928-343	Sequence 343, App
12	143	6.8	467	7 US-11-293-697-3606	Sequence 3606, App
13	140.5	6.7	411	7 US-11-293-697-4496	Sequence 4496, App
14	136.5	6.5	1460	7 US-11-247-437-14	Sequence 14, Appl
15	136	6.5	243	7 US-11-296-092-42	Sequence 42, Appl
16	133	6.3	262	6 US-10-953-349-17300	Sequence 17300, A
17	133	6.3	320	6 US-10-953-349-17299	Sequence 17299, A
18	133	6.3	502	6 US-10-511-937-3008	Sequence 3008, App
19	132	6.3	161	6 US-10-953-349-18169	Sequence 18169, A
20	132	6.3	167	6 US-10-953-349-18168	Sequence 18168, A
21	131.5	6.3	278	6 US-10-861-934-16	Sequence 16, Appl
22	131.5	6.3	278	6 US-10-861-934-26	Sequence 26, Appl
23	131.5	6.3	593	7 US-11-293-697-3668	Sequence 3668, App
24	131	6.2	250	7 US-11-296-092-76	Sequence 76, Appl
25	129.5	6.2	382	6 US-10-478-743B-2	Sequence 2, Appli

26	128.5	6.1	318	6 US-10-953-349-36014	Sequence 36014, A
27	128.5	6.1	410	6 US-10-953-349-36013	Sequence 36013, A
28	128.5	6.1	448	6 US-10-953-349-36012	Sequence 36012, A
29	128	6.1	172	6 US-10-953-349-18002	Sequence 18002, A
30	127	6.1	602	7 US-11-293-697-3416	Sequence 3416, App
31	126.5	6.0	416	7 US-11-293-697-4319	Sequence 4319, App
32	125.5	6.0	415	7 US-11-293-697-4534	Sequence 4534, App
33	125	6.0	953	7 US-11-312-958-56	Sequence 56, Appl
34	123.5	5.9	523	6 US-10-953-349-35675	Sequence 35575, A
35	123.5	5.9	783	6 US-10-511-937-2985	Sequence 2985, App
36	120.5	5.7	265	6 US-10-953-349-27984	Sequence 27984, A
37	120.5	5.7	265	6 US-10-953-349-27983	Sequence 27983, A
38	119.5	5.7	133	6 US-10-953-349-29375	Sequence 29375, A
39	119.5	5.7	213	6 US-10-953-349-29374	Sequence 29374, A
40	119.5	5.7	455	6 US-10-953-349-28615	Sequence 28615, A
41	119.5	5.7	472	6 US-10-953-349-28614	Sequence 28614, A
42	119.5	5.7	512	6 US-10-953-349-28613	Sequence 28613, A
43	119.5	5.7	663	7 US-11-293-697-2477	Sequence 2477, App
44	119	5.7	950	7 US-11-293-697-4492	Sequence 4492, App
45	118	5.6	281	6 US-10-511-937-2473	Sequence 2473, App

ALIGNMENTS

RESULT 1
 US-10-985-570-2
 ; Sequence 2, Application US/10985570
 ; Publication No. US20060100138A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Olsen, David R.
 ; APPLICANT: Polarek, James W.
 ; APPLICANT: Yang, Chunlin
 ; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
 ; FILE REFERENCE: FP0305 US
 ; CURRENT APPLICATION NUMBER: US/10/985,570
 ; CURRENT FILING DATE: 2004-11-10
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 2
 ; LENGTH: 1464
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-985-570-2

Query Match	12.7%	Score	265.5	DB	6	Length	1464
Best Local Similarity	29.7%	Pred. No.	1.8e-10				
Matches	78	Conservative	16	Mismatches	80	Indels	89
Gaps	9						
QY	13	PAAAPRGRSQG-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLAHLTLCCY	61				
DB	774	PAGAPGDKGESGSPGAGPTGARGAPGDRGEPGCPA-----GFAG-----	815				
QY	62	LELRSELRRERGAESRLGSGTPTGTLSSLGGLDPSDPSITSHLGQSPKQOPLPGEA	121				
DB	816	-----PPGAGGQCGAKGEPGAGKADGADGPPGAPG-----AGPPGPIGNVAGPA-	861				
QY	122	ALHSDSQDGHQMALNFFFPDEKPYSEBESRRVRNKRKSKSNEGADGPVKNKKKGKAGP	181				
DB	862	-----KGARGSAGPPGATGFPFGAAGRVGP	885				
QY	182	PGPNPGPPGPPGPPG-----QGPPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPGIPG	235				
DB	886	PGPSGNAGPPGPPGPPGAGKGGKGRGTGTPAGRGE--VGPPGPPGPPGAGKGGPPGADGPA	943				
QY	236	GAADKAGTFR--ENQPAVVHLOGQ	256				
DB	944	GAGTGGGAGTGGAGTGGVGLPQ	966				

RESULT 2
 US-10-985-570-3
 ; Sequence 3, Application US/10985570

; Publication No. US20060100138A1	
; GENERAL INFORMATION:	
; APPLICANT: Oleen, David R.	
; APPLICANT: Polarek, James W.	
; APPLICANT: Yang, Chunlin	
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS	
; FILE REFERENCE: FP0305 US	
; CURRENT APPLICATION NUMBER: US/10/985,570	
; CURRENT FILING DATE: 2004-11-10	
; NUMBER OF SEQ ID NOS: 3	
; SOFTWARE: PatentIn version 3.2	
; SEQ ID NO 3	
; LENGTH: 1366	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-985-570-3	
Query Match 12.5%; Score 263; DB 6; Length 1366;	
Best Local Similarity 32.9%; Pred. No. 2.5e-10;	
Matches 80; Conservative 10; Mismatches 89; Indels 64; Gaps 10;	
QY 13	PAAPRERSQCGC-----GAPARAGE-GNSCLLFLGFFGLSLALHLTLCCYLELRSE 67
DB 638	PSGLPGERGAAGTGGKGBKGFGLRGEIGNP-----GRDG----- 673
QY 68	LRRRGAESRLGSGTPTGTSGLTSSLGGLDPDPSITSHLQSPKQOPLPGEAALHSDS 127
DB 674	---ARGAHGAVAGAPGATGDRGEAGAAGPAGP-----AGPRGSPGERGEVG----- 718
QY 128	QDGHQALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKGK--KAGPPGN 185
DB 719	-----PAGNPGFAGPAGAAGQCGAKGERGKGKGENGVVGT-----GPVGAAGPAGPN 768
QY 186	GRPGPPPGPQPGPGIPGIPGTTVMGPP-----GPPGPPGPGPGLQGPSGAAD 239
DB 769	GPPGPAAGSGDGGPPGQMTGFPGAAGRT--GPPGPSGISGPPGPGAGKEGLRGRDQG 826
QY 240	KAG 242
DB 827	PVG 829
RESULT 3	
US-10-985-570-1	
; Sequence 1, Application US/10985570	
; Publication No. US20060100138A1	
; GENERAL INFORMATION:	
; APPLICANT: Oleen, David R.	
; APPLICANT: Polarek, James W.	
; APPLICANT: Yang, Chunlin	
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS	
; FILE REFERENCE: FP0305 US	
; CURRENT APPLICATION NUMBER: US/10/985,570	
; CURRENT FILING DATE: 2004-11-10	
; NUMBER OF SEQ ID NOS: 3	
; SOFTWARE: PatentIn version 3.2	
; SEQ ID NO 1	
; LENGTH: 1466	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-985-570-1	
Query Match 12.2%; Score 256; DB 6; Length 1466;	
Best Local Similarity 28.8%; Pred. No. 7.8e-10;	
Matches 86; Conservative 12; Mismatches 97; Indels 104; Gaps 10;	
QY 13	PAAPRERSQG-----CGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCCYLELRSEL 68
DB 793	PRGSPGERGETGPPGAPGAPGAPGQNGEPGG----- 823
QY 69	RRRGAESRLGSGTPTGTSGLTSSLGGLDPDPSITSHLQSPKQOPLPGEAALHSDQ 128
DB 824	KGERGAPGKEGGPGVAGTGGGFPAGPGP-----QGVKGERGSPGPGGAAG----- 873
RESULT 4	
US-10-505-928-493	
; Sequence 493, Application US/10505928	
; Publication No. US20060088532A1	
; GENERAL INFORMATION:	
; APPLICANT: Ludwig Institute for Cancer Research et al.	
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES	
; FILE REFERENCE: 28967/39178	
; CURRENT APPLICATION NUMBER: US/10/505,928	
; CURRENT FILING DATE: 2004-08-27	
; PRIOR APPLICATION NUMBER: US 60/363,019	
; PRIOR FILING DATE: 2002-03-07	
; NUMBER OF SEQ ID NOS: 866	
; SOFTWARE: PatentIn 3.2	
; SEQ ID NO 493	
; LENGTH: 744	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-10-505-928-493	
Query Match 12.0%; Score 252; DB 6; Length 744;	
Best Local Similarity 25.6%; Pred. No. 7.3e-10;	
Matches 112; Conservative 26; Mismatches 145; Indels 154; Gaps 17;	
QY 2	GYPEVERRELLPAAAPRERSGOGCGGAPARAGSGNSCLLFLGFFGLSLALHLTLCCY 61
DB 328	GFPGKGEGGLP-GLPFGPGLGIGKPGFPFGKGD-----RGMGVFGAL----- 371
QY 62	LELRSELRERRGAESRLGSGTPTGTSGLTSSLGGLDPDPSITSHLQSPSK-----Q 113
DB 372	-----GPRGKGPICAGGIGGPPGPGPLGIPGPGWGPCAI----GFGPKGEGGIVGPG 422
QY 114	QPLEP-GEAALHSDSDQGHQALLNFFPDPEKPYSEESRRVRNRKSKSNEGADGPVN 172
DB 423	GPPGKGBPGLQG-----PPGKPGFLGEVGPQGM-----GLPGPIGP 460
QY 173	K-KKGKKA-----GPP-----GPNPGPPGPPGPPGPPG 199
DB 461	KGEAGKGVPLGVLGVLGPKGERGIPGDQGLQGPPIGIGGSGPIGPIGPIGPKGE 520
QY 200	PGIPGIPG-----TTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 253
DB 521	PGLPGPPGPPGPGKPGVAGLHGPPCKPGALGPGQGLGPPGPPGPPGPPGPPGPPG 580
QY 254	QGQ-----GSAIQVNKDLGGVLNDSRITMN-----PKVF----- 284
DB 581	QGEYLPDMGLGIDGVKPPHAYGAKKGKNGGAYEMPAFTAELTAPFPVPVGFKNKLY 640
QY 285	----KLHPSRSELEVLVDGTYFIYSQV-----EVYINFTDFA 318
DB 641	NGRQNYNFGTIGFTCEVPGVYFAFVHVCKGNGVWVAFKQNEPVMYTYDEYKKGFLDQA 700
QY 319	SYEVVVDEKP-----ELQ 331
DB 701	SGSAVILLRPGDRVFLQ 717
RESULT 5	
US-10-505-928-438	

